FEBRUARY
1954
VOLUME 18
NUMBER 1

JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

JOURNAL OF CONSULTING PSYCHOLOGY

Edited by Laurance F. Shaffer

Teachers College, Columbia University

Associate Editors: Edgar A. Doll, *Bellingham, Wash.* Ann Magaret Garner, *University of Illinois.* William A. Hunt, *Northwestern University.* E. Lowell Kelly, *University of Michigan.* Morris Krugman, *Board of Education, New York, N. Y.* Bertha M. Luckey, *Board of Education, Cleveland, Ohio.* Fred McKinney, *University of Missouri.* Anne Roe, *New York, N. Y.* Carl R. Rogers, *University of Chicago.* R. Nevitt Sanford, *University of California.* Edward Joseph Shoben, Jr., *Teachers College, Columbia University.* J. R. Wittenborn, *Yale University.*

Lorraine Bouthilet, *Managing Editor*

The *Journal of Consulting Psychology* is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy and counseling, personality, and the dynamics of behavior. Case studies, theoretical contributions, descriptions of clinical techniques, and discussions of the professional problems of clinical psychology also appear in its pages.

Editorial Office

Manuscripts submitted for publication, books and tests for review, and correspondence on editorial matters should be sent to the Editor, Laurance F. Shaffer, 525 West 120th St., New York 27, N. Y.

Information for contributors: The important characteristics of a good article are the significance of its subject matter, the demonstrated soundness of its conclusions, its clarity, and its brevity. Clarity is improved by careful outlining, discriminating use of headings, and a simple, straightforward style of writing. For brevity, authors should strive to convey every thought without waste of words, and should avoid repetitions of methods, data, or findings.

The style of all manuscripts should conform to the instructions given in the Publication Manual of the American Psychological Association, prepared by the Council of Editors (*Psychological Bulletin*, 1952, 49, 389-449). If it is not otherwise available, the *Manual* may be purchased, for $1.00, from the Washington office of the Association. The Editor expects each author to study the *Manual* carefully before submitting a manuscript. Special attention should be given to the style of headings, capitalization, indention, spacing, and footnotes, and to the preparation of tables, figures, and lists of references. Otherwise acceptable manuscripts which depart seriously from the standards of the *Manual* are returned to the author for revision and retyping.

There is a lag of about nine to twelve months between the acceptance of an article and its publication. Authors may secure immediate publication, however, by arranging for articles to be printed as extra pages, increasing the number of pages that subscribers receive. The cost to the author of such early publication is approximately $20.00 per printed page.

Authors receive fifty reprints gratis, and may order additional quantities when returning proof.

Business Office

Communications concerning subscriptions, change of address, claims for the nonreceipt of a number, advertising, and other business matters, should be sent to the American Psychological Association, Inc., 1333 Sixteenth St., N. W., Washington 6, D. C.

The Journal of Consulting Psychology is published bimonthly in February, April, June, August, October, and December, the yearly volume comprising approximately 480 pages. The subscription per year is $7.00, foreign $7.50, single number $1.25.

Copyright, 1954 by the American Psychological Association, Inc. Printed and mailed by the Dentan Printing Co., 23 West Colorado Ave., Colorado Springs, Colo. and entered as second-class matter at the Post Office at Colorado Springs, Colo. Acceptance for mailing at special rate of postage provided for in section 538, Act of February 25, 1925, authorized May 27, 1940.

JOURNAL OF CONSULTING PSYCHOLOGY

February, 1954 Vol. 18, No. 1

Some Behavioral Correlates of Rorschach's Experience-Type: *Jerome L. Singer and Herbert E. Spohn* - - - - - - - - - - - - - 1

A Comparison of Raven's Progressive Matrices (1938) with a Short Form of the Wechsler-Bellevue: *Bert Levine and Ira Iscoe* - - - - - - - - - - 10

Rorschach Performance in Relation to Two Types of Stress: *Murray S. Stopol* - - - 11

Age, Sex, and the Manifest Anxiety Test: *A. W. Bendig* - - - - - - - - - - 16

The Rorschach White Space Response and "Oppositional" Behavior: *Albert Bandura* - 17

Acceptance of Self as Measured by Interviews and the Index of Adjustment and Values: *Robert E. Bills* - - - - - - - - - - - - - - 22

Prediction of Aggression from the Rorschach: *Winifred Ingram* - - - - - - - - 23

The Effects of Figure-Ground Contrast Upon Perception as Evaluated by a Modified Rorschach Technique: *Arvord W. Belden and E. Earl Baughman* - 29

Interrelationships among Several Measures of Anxiety and Hostility: *Leonard D. Goodstein* - - - - - - - - - - - - - - - - - - - 35

Mosaic Productions in Cerebro-arteriosclerosis: *Brendan A. Maher and Anthony W. Martin* - - - - - - - - - - - - - - - - - - 40

Sexual Symbolism: A Validity Study: *Leon H. Levy* - - - - - - - - - - - - 43

The Interpretations of Anxiety by Various Psychotherapeutic Schools: *John H. De Michele* - - - - - - - - - - - - - - - - - - - 47

The Relationships among Projection, Empathy, Reality, and Adjustment, Operationally Defined: *Ralph D. Norman and Patricia Ainsworth* - - - - - - 53

Moral Indignation: A Correlate of Denied Dislike of Parents' Traits: *Sidney M. Jourard* 59

The Comparability of Measures of Profile Similarity: An Empirical Study: *James N. Mosel and June B. Roberts* - - - - - - - - - - - - - 61

Children's Responses to Animal and Human Figures in Thematic Pictures: *Mary Ann Smith Armstrong* - - - - - - - - - - - - - - - - - - 67

Delinquent vs. Nondelinquent Performance on the Porteus Qualitative Maze Test: *Richard F. Docter and C. L. Winder* - - - - - - - - - - - - 71

New Books and Tests - 74

Brief Reports:
A New Publication Procedure

For a trial period, the *Journal of Consulting Psychology* will accept *"brief reports"* for early publication without expense to the author, under certain conditions. The new procedure is intended to reduce publication lag and permit the circulation of worth-while studies of specialized interest or minor importance which cannot be accepted at present because of limitations of space.

A *brief report,* strictly limited to one printed page, may be accepted if the author agrees to make a more extended report available. Up to four pages in each issue may be devoted to brief reports, published in the order of their receipt without regard to the dates of receipt of articles accepted for regular publication. It is anticipated that most of the brief reports can appear in the first or second issue to go to press following their acceptance.

The procedure for submitting a *brief report* is:

1. The author sends the editor *both* the brief report and a more complete report of the study. The full report should be of sufficient length to give a clear account of the background, procedure, results, and conclusions of the study.

2. The *full report* is prepared in the style specified by the *Publication Manual* [1], except that it may be typed with single spacing. The editor will send the full report to the ADI.

3. The *brief report* gives a clear, condensed summary of the procedure of the study, and as full an account of the results as space permits. To insure that the brief report does not exceed one printed page, the author prepares it according to these specifications:

The text of the brief report, including all matter except the title and the author's lines, does not exceed 80 typewritten lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, or 4.2 inches long in pica. The manuscript is double spaced throughout and, except for the short lines, follows the standard style [1].

Headings, tables, and references are avoided in the brief report. If essential, they must be counted in the 80 lines. The brief report includes, within its 80 lines, a footnote in this style:[1]

[1]An extended report of this study may be obtained without charge from John Doe, 300 Market St., Prospect 6, Mass. [giving the author's full name and address] or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. .... from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $..... for microfilm or $..... for photocopies.

4. The author of a brief report prepares at least 100 mimeographed copies of the extended report. He agrees: (*a*) to send copies of the extended report upon request as long as the supply lasts, and (*b*) not to submit the extended report for publication to another printed journal of general circulation.

Reference

1. American Psychological Association. Council of Editors. Publication manual of the American Psychological Association. *Psychol. Bull.,* 1952, 49, 389-449.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Some Behavioral Correlates of Rorschach's Experience-Type

Jerome L. Singer
Franklin D. Roosevelt VA Hospital, Montrose, N. Y.

and Herbert E. Spohn
Teachers College, Columbia University

Within the framework of the Rorschach method the experience-type, the ratio of human movement (*M*) to weighted color responses (*Sum C*), although central to Rorschach's thinking [15], is in actual practice rather difficult to employ, as Beck [1] has pointed out. Part of the difficulty stems perhaps from Rorschach's somewhat arbitrary juxtaposition of two types of response which result from rather disparate perceptual processes. The situation is ameliorated somewhat by the fact that Rorschach has presented fairly specific indications of the behavioral significance of each of the two modes of response as well as of the overt behavior occurring in persons with a clear preponderance of one component of the ratio. As part of a program of research designed to integrate key Rorschach determinants with general theories of perception and thinking, one focus of this study will be on the correlates of Rorschach's experience-type as given in the *M*:*Sum C* ratio. Because of the interesting theoretical implications of the human movement response, however, primary emphasis in this study will be upon the *M* component of Rorschach's ratio. The color responses will be considered primarily from the point of view of how their presence in a record modifies the observed relationship between *M* and overt behavior.

Largely on the basis of intuition and empirical observation Rorschach [15] proposed that a triadic relationship exists between overt motor activity, perception of movement, and imaginative processes. He indicated that perception of humans in motion on the inkblots was correlated on the one hand with "greater creative ability," "more 'inner' life," and on the other hand, with "stable, measured motility" [15, p. 78]. He specifically asserted that "Kinesthetic engrams, therefore, act as inhibitors of physical activity; motor activity inhibits kinesthetic engrams" [15, p. 80]. Schachtel has summarized this view of the kinesthetic (*M*) response very effectively in his thoughtful paper on projection. He writes:

> . . . the kinesthetic percept is accompanied by arrested, checked motor innervations. Instead of being carried out, these innervations lead to empathy with movement observed in the environment, or to kinesthetic imagery as in kinesthetic percepts in Rorschach's test or as in dreams where motor activity is suspended by sleep [16, p. 87].

The challenge of this triadic relationship lies in its ready application to theories and observations concerning the development of thought which stem from diverse investigators in the fields of psychology and psychiatry. Thus, Freud in presenting his concept of the *reality-principle* as a key phase of ego development also pointed to the relationship between thought and restraint of motor discharge. He described thought as "endowed with qualities which made it possible for the mental apparatus to support increased tension during a delay in the process of discharge. It is essentially an experimental way of acting . . ." [4, p. 16]. Other psychoanalytic theorists [2, 3, 13] have elaborated on Freud's views but generally seem to concur in relating the origin of thought to a delay or inhibition of immediate

[1]Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions do not necessarily reflect the opinion or policy of the Veterans Administration. The authors are indebted to clinical psychologist trainees Jack Herman and Paul Perez as well as to the staff of the Psychology Service for assistance in various phases of the investigation.

need gratification through motor activity.

Somewhat similar theories supported with some empirical data are presented by Piaget [11], Lewin [8], and Murphy [10]. All have in common a conception that the development of imagination and planful fantasy are concomitants of the increased restriction of the growing child's motility as well as of the enforced socialization of his egocentric speech patterns. To the extent that Rorschach's *M* represents fantasy activity, developmental findings with the inkblots are consistent with these theories. Thus, *M* first appears as a rule in children's records at age six, at a time when school attendance markedly curbs motility and enforces socialization [7, 19]. Differences between schizoid and normal children [19] or between disturbed children producing different amounts of *M* [7] also appear to be in accord with these theories.

With respect to the actual perceptual process in the production of Rorschach movement responses, experimental literature in perception and physiological psychology offers some clues [20, 21, 23]. Werner [21], for example, has hypothesized that a common dynamic property, *tonicity,* underlies both sensory and motor responses. Available tonic energy may be released through body movement or may increase tonicity in a sensory area bringing about spatial displacement and the illusion of motion. Experimental support for this point of view with particular implications for Rorschach theory has come in a series of experimental studies [5, 23] and in comparisons of the perceptual and Rorschach behavior of feebleminded children who represent contrasting tendencies to motor activity [21, 22]. Thus, the hyperactive exogenous feebleminded showed fewer Rorschach movement responses than did the motorically phlegmatic endogenous feebleminded [22]. Evidence that the number of *M* produced on the inkblots by normals increases following a period of motor inhibition [9, 18] also supports Werner's hypothesis. Of particular interest for the present study was the finding that the number of *M* initially produced by subjects was positively correlated with their subsequent ability to inhibit motor response [9].

Concern in the present investigation will be with *M* in its hypothetical relation to motor behavior rather than with its presumptive correlates in fantasy and imagination. To ascertain the generality of findings obtained previously with normals, this study was extended to include a psychopathological group, i.e., schizophrenic males. In addition, a situation designed to elicit evidence of spontaneous motor activity and environmental responsiveness was utilized, since previous research [9, 18] had been limited to fairly specific motor inhibition tasks. It was felt that on theoretical grounds the presence of *M* responses in the Rorschach records of schizophrenics would bear the same general relation to motor inhibition that was observed in studies with normals. In the sphere of spontaneous motor activity, however, the color responses in the *M*: *Sum C* ratio might be expected to play a role in determining the degree of responsiveness. This study, moreover, involves the behavioral correlates of Rorschach's *M* and *C* categories as elicited in standard clinical administrations of the test rather than in the quantitative changes in these responses following inhibition as in the earlier work [9, 18]. For this reason the present study may be expected to throw some light more directly upon clinical interpretations of the response categories oriented toward description of personality structure.

The Experiment

Hypotheses. The following hypotheses were formulated:

1. Schizophrenic groups chosen as extremes of introversive or extratensive experience-types will differ significantly in two behavioral situations. Introversives ($M > C$) will show longer inhibition times in a motor inhibition task and less motor activity during a waiting period than extratensives ($C > M$).

2. Schizophrenic *S*s who have produced several human movement responses in their Rorschach protocols (High-*M* group) will show longer inhibition times and less waiting-room activity than an equated schizophrenic group who have produced relatively few Rorschach *M* (Low-*M* group).

3. The greater the number of Rorschach color responses in the patients' records, the greater will be the tendency for environmental responsiveness through motor activity during the waiting period. If, however, Hypothesis 2

is confirmed, the quantity of color responses may be expected to exercise a modifying rather than a determining influence upon the amount of motor activity.

Subjects. The *S*s were 50 male veteran patients officially diagnosed as schizophrenic. They were divided into two groups of 25 *S*s each on the basis of the *M* responses in their Rorschach protocols.[2] Those producing more than one *M* were designated as the High-*M* group while those whose records yielded one *M* (usually Card III) or less composed the Low-*M* group. These groups were equated for other variables which might conceivably be related to the experimental situation, e.g., age, duration of hospitalization, IQ, and Rorschach scores. Table 1 presents comparisons of the means for all groups on certain pertinent variables. Others which showed low incidence are omitted as are percentages. For the High- and Low-*M* groups only the difference in *M* was significant. Both groups contained approximately equal numbers of the various subtypes of schizophrenia, predominantly paranoids and hebephrenics in this sample.[3]

In order to test the hypothesis concerning introversion and extratensiveness, two equated groups of *S*s chosen as the extremes of the *M*: *Sum C* ratio for the patient population were also employed. These groups were chosen to exclude *S*s with "coarctated" or "ambiequal" ratios. These groups consisted of 19 introversive *S*s and 16 extratensives who were equated on all variables except *M* and *C*. Data for these groups are also presented in Table 1. Only the differences in *M*, *CF*, *C*, and *Sum C* are significant.

Procedure. The *S*s underwent the following sequence of procedures:

1. Waiting-room observation

[2]It should be noted that this study is limited to the role of the human movement determinant with all other movement percepts, animal and inanimate movement (*FM* and *m*), controlled. Previous studies [9, 18, 22] have yielded mixed results, occasionally suggesting that *FM*, *m*, and *M* have common behavioral correlates, occasionally suggesting differences between them. Since Rorschach's original formulations were clearly related to the experience of *human* movement, the *M* responses have been singled out in this study for the sake of simplicity in design.

[3]Diagnoses were made by admitting psychiatrists and were corroborated by ward psychiatrists. There seemed to be little relationship between the subtype of schizophrenia and the Rorschach protocols of the 200 patients from whom the experimental samples were drawn. Most patients revealed paranoid delusions after some probing in any case, and patients with rather similar delusional content could be found in both High- and Low-*M* groups.

Table 1

Group Means for High- and Low-*M* and Introversive-Extratensive Groups on Age, IQ, and Pertinent Rorschach Variables

Matching variables	Groups			
	High-*M* ($N = 25$)	Low-*M* ($N = 25$)	Introversive-extratensive ($N = 19$)	($N = 16$)
Personal data				
Age	32.56	32.88	33.84	32.19
IQ (Wechsler-Bellevue)	116.50	113.00	114.00	120.78
Rorschach determinants				
R	24.84	19.36	22.00	22.94
M	3.84*	.60*	4.05*	1.13*
FM	4.16	3.04	4.11	3.19
m	.60	.52	.68	.50
C'	.80	1.16	.53	1.56
c	1.80	.84	1.32	.94
F	10.08	8.76	9.05	9.81
F + % (Beck)	83.88	80.60	86.53	82.13
FC	1.72	1.88	1.20*	2.31*
CF	1.04	1.88	.42*	2.81*
C	.48	.32	.08*	.94*
Sum C	2.60	3.22	.92*	5.34*

* Differences between the means of High- and Low-M groups in M and between means of introversive and extratensive groups in M, FC, CF, C, and Sum C are significant beyond the .05 level.

2. Motor inhibition task
3. Brief interview

The waiting-room observation was a procedure modified from studies carried out by Rickers-Ovsiankina [14]. The *S* was conducted into the ward visitors' room, a customary waiting room in the hospital. He was informed that *E* would call for him in about 15 minutes and was told to spend his time as he chose. The *S* was seated near a table on which were spread a large number of puzzles and colorful test objects, including such things as Isochromatic Color Plates, Lowenfeld Mosaic pieces, various mechanical aptitude tests such as nut and bolt assembly, etc. At the far end of the room, apparently working on some papers, was the observer, who was almost always unknown to the patient. The observer recorded, according to prearranged plan, all movements, activity, and speech of the patient during the 15-minute period. Recording sheets broken down into 15-second intervals were used, and the observer entered each change in position, obvious shift in attention, gross movement or comments, manipulation of materials, etc. These observations proved relatively easy to make, particularly since there were no others in the room and clear vision of *S*'s activity was possible. The presence of the observer did not seem to disturb the patients, who were accustomed to finding strangers seated in this room while waiting for psychiatric consultations. Subsequent questioning revealed that only two paranoid patients guessed the function of the observer. Any comments addressed to the observer by *S* were included in the protocol as part of a response to environmental stimulation.

Protocols were rated on a six-point scale for degree of motor activity and environmental response. Each point was defined as an aid in rating based upon a preliminary sampling of the various activities manifested by patients in a waiting-room situation. The definitions were as follows:

0 = No gross motor activity other than stretching, scratching, smoking, minor finger movements, foot tapping.

1 = No gross motor activity other than playing with personal objects, e.g., glasses, shuffling cards, toying with pencils, etc.

2 = Some minimal movement around room, reading, looking at objects on table (without manipulating them), or doing crossword puzzle brought into the room by *S*.

3 = Some speech, some movement around room, evidence of hallucination, laughing, loud and persistent whistling, other gross motor activity, e.g., picking up one or two objects from table and examining them.

4 = Fairly extensive motor activity, playing with many objects on table, or a few such objects for a very long time, occasional comments to observer or to self.

5 = Considerable restless motor activity or speech, noticing and manipulating many objects, addressing observer or persons passing outside the room.

Protocols were rated by two judges without knowledge of *S*'s name or group. One judge made separate ratings a month apart as well. Rater reliability was determined by product-moment correlations. Final waiting-room data consisted of the average of the three ratings.

The motor inhibition task consisted of a procedure adapted from the Downey Will-Temperament Scale. This task called for *S* to write the phrase "New Jersey Chamber of Commerce" as slowly as he possibly could without lifting pencil point from the page or stopping motion of the pencil. This task had been employed in prior research in this series [9]. The *S*s were informed that they were helping in the standardization of a new testing procedure. The task proved relatively simple and was easily comprehended by the patients, who cooperated readily enough. As a check on cooperation a writing sample at normal tempo was obtained to be sure that *S* had made a definite effort to inhibit his motor response. The test phrase was written in pencil across an 8-inch page between ½-inch parallel lines to maintain approximate uniformity in size. Data for this task were based on the inhibition times taken in seconds.

The interview followed immediately and consisted of an exploration, within the limits of the patient's accessibility, of *S*'s daydream and other thinking activity, its preferred form and content. Patients' comments concerning their own motor activity were considered as well. It was not possible at this time, however, to get very much information concerning motivational aspects of the patients' fantasy life.

The Rorschachs, administered originally by the hospital staff in the course of diagnostic examinations, were rescored by the senior author. In the few cases of discrepant scoring the judgment of a third experienced clinician was relied upon for a final decision. Klopfer's scoring was used.

Results

The human movement determinant. The relative simplicity of the motor inhibition situation resulted in good cooperation by patients.[4] Even those *S*s with lowest inhibition times wrote the test phrase at least four times as slowly as the normal writing sample. Scoring of the waiting-room protocols according to the activity scale also proved relatively simple. The average interrater reliability coefficient (based on *z* transformations) was $r = +.85$. Data

[4] A third experimental situation, mirror tracing, had to be abandoned early in the development of the study when it proved too difficult and disconcerting to *S*s, who frequently rejected it.

presented for this variable are based on the average of three ratings. Results for the High- and Low-*M* groups and the introversive-extratensive group on both variables are presented in Tables 2 and 3. It should be noted that where specific directional differences were predicted the value of $p/2$ is employed to indicate significance.[5]

Table 2
Inhibition Times and Waiting-Room Activity Scores for High– and Low–*M* groups

Groups	Experimental situation means	
	Inhibition time (secs.)	Activity rating
High–*M*	277.56**	1.84**
Low–*M*	138.40	3.09

** Difference between group means is sig. at $p < .01$.

Table 3
Inhibition Times and Waiting-Room Activity Ratings for Introversive and Extratensive Groups

Group	Experimental situation means	
	Inhibition time (secs.)	Activity rating
Introversives ($M > Sum\ C$)	266.16**	1.66**
Extratensives ($Sum\ C > M$)	139.63	3.28

** Difference between group means significant at $p < .01$.

From the data it seems clear that the High-*M* group shows considerably longer inhibition time and considerably lower activity scale ratings than the Low-*M* group. These results tend to support the hypotheses. A product-moment correlation between *M* and inhibition time yields an *r* of .29, which is significant at $p < .05$. The product-moment *r* for *M* and waiting-room activity scores is −.23 which barely misses significance. The correlation between *M* and inhibition time is somewhat lower than that obtained in earlier studies with normal *S*s [9], but this most likely reflects the typically greater variability of schizophrenic *S*s as well as some procedural differences between the experiments.

[5]Since the contrasted groups did not differ significantly in total responses (*R*) the absolute number of *M* was employed. Statistical checks on the influence of *M*% and other determinants as a function of *R* did not alter the trend of the data.

Previous studies had suggested the value of considering certain qualitative aspects of the *M* response [18, 19]. Certainly a corollary of the major hypothesis would be that those *S*s who produce *M* responses that are more active or vigorous in the ascribed movements may be expected to show greater inhibition and less spontaneous motor activity in their own behavior. To test this possibility the human movement responses for the 25 High-*M* *S*s were rated for *active* ("Man running, girls dancing") versus *static* movement ("Woman staring, man sleeping"). It was possible then to subdivide the High-*M* group into the 15 *S*s who showed two or more active *M* and the 10 who showed less than two such responses. The results presented in Table 4 reveal that even within the High-*M* group significant differences emerge; those *S*s who emphasize active *M* tend to have longer inhibition times and lower activity scores than the patients who gave more static *M*. When the total sample of 50 *S*s was broken down into those with preponderantly active or static *M* a similar finding emerged. These rather clear-cut findings lend further support to the hypothesized relationship between motion perception and motor inhibition. They also suggest, pending further research, a necessary refinement in interpretation of Rorschach, i.e., consideration of the vigorousness of movement. Exploratory scoring of *M* as flector and extensor and aggressive ("men with clubs") or passive ("young cherubs dancing") content failed to yield any additional information relevant to the hypotheses.

The experience-type. Results in motor inhi-

Table 4
Inhibition Times and Waiting-Room Activity Ratings for High–*M* *S*s Producing Active and Static *M*

Group	Experimental situation means	
	Inhibition time (secs.)	Activity rating
Active-*M* ($N = 15$)	377.9**	1.40*
Static-*M* ($N = 10$)	127.1	2.50

* Difference between group means significant at $p < .02 > .01$.

** Difference between group means significant at $p < .01$.

bition and waiting-room activity for the introversive and extratensive experience-type groups are presented in Table 3. They reveal significant differences paralleling those with *M* alone. Introversive *S*s show significantly longer inhibition times and significantly less waiting-room activity than do extratensives. Since these results, obtained with groups differing in both *M* and *C*, are quite similar to those obtained when only the number of *M* was used in separating groups, they suggest that it is the *M* response alone rather than the *M*:*Sum C* ratio which is crucial in relation to motor activity. It should be kept in mind that several *S*s in the High-*M* group were technically extratensive; they showed even more color responses than human movement responses. This dominant role, played by the *M* response alone, while in accord with the sensory-tonic theory casts some doubt upon Rorschach's emphasis on the experience-type per se.

Closer scrutiny of the results, however, does indicate the possibility that the color responses in the ratio do have some influence, particularly in the waiting-room situation, which may evoke spontaneous motor response. Thus, inspection of Table 2 indicates that when *M* alone is the basis for group division the difference in waiting-room scores for the contrasting groups is somewhat below that which obtains when both *M* and *C* are considered as in Table 3. In the breakdown by active-static *M* in Table 4, the waiting-room score differences between the groups are again lower, although significant. In both High- and Low-*M* and introversive-extratensive analyses, partialling out of the color factor reveals a slight although insignificant increase in the negative correlation of *M* and waiting-room activity.

Far more suggestive of a real influence of color, however, are the results obtained when the actual ratio of *M*:*Sum C* rather than introversive-extratensive types are considered. Statistical treatment of this ratio involved division of the High- and Low-*M* *S*s into High- and Low-*C* subgroups using a *Sum C* of 3.0 (the sample mean) or less as a basis for assignment to the Low-*C* group. Results are presented in Table 5, which reveals that motor inhibition scores vary chiefly as a function of the number of *M*. On the other hand, a marked increase

Table 5

Inhibition Times and Waiting-Room Activity Ratings for High- and Low-*M* *S*s Subdivided by *Sum C* Scores

Group	Rorschach means		Experimental situation means	
	M	*Sum C*	Inhibition time (secs.)	Activity rating
High *M*: Low *C* ($N = 17$)	3.1	0.8	247.3	1.66
High *M*: High *C* ($N = 8$)	5.1	6.3	298.0	2.23
Low *M*: Low *C* ($N = 13$)	0.4	1.4	139.4	2.82
Low *M*: High *C* ($N = 12$)	0.8	5.0	127.0	3.34

in waiting-room activity scores is observable as we move from the High-*M*:Low-*C* group through High-*M*:High-*C*, Low-*M*:Low-*C*, and, finally, Low-*M*:High-*C*. This is so despite the fact that the High-*M*:High-*C* group actually averaged more *M* responses than the High-*M*:Low-*C* group. Although the differences from subgroup to subgroup are not significant, the regularity of the increase in activity with the increasing weight of color in the experience-type ratio points to a clear influence of the color factor upon the waiting-room behavior. Analysis of variance for the waiting-room activity scores in the four subgroups yields an *F* of 4.36 which is significant beyond the .01 level.

Qualitative observations. Although each patient was interviewed briefly following the motor inhibition task, the severity of the pathology in many cases prevented satisfactory communication. To some extent those individuals who reported tendencies toward active fantasy lives such as heroic daydreams or reviewing of past events or future plans tended to fall in the High-*M* group. Some *S*s who showed Low-*M* production but comparatively little motor activity reported that they tended to lose themselves in viewing television or in focusing on minute bodily changes with consequent somatic delusions.

On the whole, patients in the Low-*M*:High-*C* category tended to have spent more time on the "violent" ward because of hyperactive behavior. High-*M*:Low-*C* patients were more

likely to be found on the privileged ward or were considered relatively "well-behaved," albeit severely ill, by ward personnel. In rough confirmation of Rorschach's [15] discussion of persons with "dilated" experience-types, it was observed that High-*M*:High-*C* *S*s tended to be among the most "interesting" patients in the hospital. They were usually better known to personnel and tended to demand and receive more attention while "coarctated" Low-*M*: Low-*C* *S*s, for example, tended to be less known by name and more inconspicuous.

An example of a patient whose record showed a High-*M*: High *Sum C* ratio and who typified characteristics of a "dilated" experience-type may be cited. This man was a severely disturbed paranoid schizophrenic whose ratio was 8:24, revealing more *M* and more *C* than any other patient in the population of 200 from which the equated groups were drawn. Although stemming from a low socioeconomic background this man attracted considerable attention because of the variety of his talents, fluency and eloquence of his speech, and his physical activity and vigor. On the one hand he revealed superior linguistic gifts and artistic talent while at the same time he was the outstanding athlete in the hospital. This man's drive for physical activity was manifested in his performing calisthenics between innings of pitching softball. He spoke rapidly, loudly, and almost interminably in group therapy or in conversation. Yet he was capable of sitting on the ward for long periods of time, almost completely motionless, weaving elaborate and highly imaginative fantasies of heroic achievements or of the plots devised against him. Both the *M* and *C* aspects of his personality could be readily observed although the hyperactivity predominated in behavior as in the experience-type ratio.

Discussion

The results of this study tend in general to support the experimental hypotheses. In doing so they suggest several implications that merit consideration theoretically as well as in clinical practice.

1. These findings support those of previous studies [9, 18, 22] in suggesting some validity for Rorschach's empirical association of the perception of humans in motion on the inkblots and tendencies toward inhibited or controlled motor response. They offer some evidence for the generality of the relationship inasmuch as they were obtained from a markedly pathological group of *S*s, schizophrenic patients, while the previous studies involved normal adults [9, 18] or mentally defective children [22]. Thus, despite the anticipated variability of the patients and the consequently lower correlation between *M* and the motor inhibition task, the trend for the production of several Rorschach *M* to be associated with comparatively greater capacity to inhibit writing speed emerges again. Although the low correlations preclude prediction in the individual case they suggest that the *M* response has a common interpretive significance across normal and abnormal groups. This does not deny the necessity for considering other factors in the Rorschach protocol which may counterbalance the *M* or alter its interpretation somewhat. At the same time it seems clear that as a perceptual or apperceptual phenomenon presumably exemplifying some general behavior principle, the human movement response to the Rorschach inkblots has definite correlates in behavior quite apart from the immediate testing situation.

2. Of particular theoretical interest is the nature of these correlates of *M*. The fact that individuals chosen on the basis of contrasting tendencies to perceive humans in motion on inkblots tend to differ in their capacity to inhibit motion or to move about in a waiting room supports the view that perceptual response reflects pervasive behavior characteristics. More specifically, theories implying a vicarious relationship between motion perception and overt movement, of which the sensory-tonic theory [21, 23] is an example, gain support from these findings. Whether Rorschach's *M* reflects imaginativeness or fantasy processes as well as tendencies toward motor restraint remains a question for further research. The positive results of this study offer encouragement for an attempt to explore the interrelations of motion perception, fantasy, and motor activity with the ultimate aim of developing the unified theory of thinking now only speculatively suggested by the writers cited in the introduction [8, 10, 11, 13].

3. The data of the experiment throw comparatively little light on the knotty problem of the experience-type ratio, however. Most of the data suggest that the quantity of *M* rather than the experience-type as given in $M > C$ vs. $C > M$ is most potent in determining mo-

tor responsiveness. The results presented in Table 5 suggest, however, that the role of color cannot be ignored. The amount of spontaneous motor activity in the waiting-room situation varied with a quantitative relationship between *M* and *C* somewhat different from that implied by introversive-extratensive ratios. The observed relationship suggests that while *M* is the primary factor related to motor inhibition, color, reflecting the individual responsiveness to the environment, to some extent counterbalances this tendency. Since the influence of color emerged only in the waiting-room situation (which might be considered evocative of more spontaneous performance), the interpretation of the color factor as reflecting environmental responsiveness or outgoing tendencies [6, 12, 16] would appear consistent with the data. Whether it is the color qua color or merely a tendency to respond in diffuse associations with lessened consideration of blot contour [17] which reflects *S*'s characteristic emotional or motor lability is a question beyond the scope of this investigation. Nevertheless, the finding that some influence of the actual experience-type ratio exists is quite important for validation of Rorschach's original interpretations and merits more intensive scrutiny.

A final word may be said concerning the implications of the results for research in the field of projective techniques. It seems clear that experimentation in the theory and process of clinical instruments is possible. The question seems to be one of finding concrete situations in which the somewhat nebulous personality descriptions can be expressed operationally. The interaction of various factors and the necessarily ideographic approach of the clinician need not constitute an insurmountable obstacle. The original assignment of interpretative meaning to the determinants on Rorschach's part was, as he makes quite clear, drawn from his observation of their correlates in overt behavior. Even in "global" interpretation some aspect of the "nuclear definition" [12] of the factor is necessarily preserved. Thus, one need not expect to find very high correlations in group data, but low, significant correlations should emerge, as they have in the present series of studies.

Summary and Conclusions

This study was part of a program of research designed to integrate key Rorschach determinants within a more general theory of perception. Previous studies have revealed a relationship between inhibition of motor response and the tendency to produce human movement percepts (*M*) on the Rorschach, a finding in keeping with Werner's sensory-tonic theory of perception [21] as well as with certain theories of thought and its development [4, 8, 10, 11, 13]. The present study extended this research to include a pathological group (schizophrenics) and also focused attention on the experience-type (*M*:*Sum Color*) ratio as well as *M* alone. It was hypothesized that schizophrenics producing several *M* percepts would show greater tendency to inhibit overt motor response than a comparable group which produced few *M*. Similarly schizophrenics with introversive experience-types were expected to show more motor inhibition and less spontaneous gross motor activity during a waiting period than extratensive *S*s. It was anticipated that the number of *M* responses would prove most effective as an index of motor inhibition, with *C* responses, however, influencing the tendency towards spontaneous motor activity.

Fifty hospitalized male schizophrenics were divided into two groups on the basis of production of two or more, or one or less, Rorschach *M* responses. These groups did not differ significantly in age, intelligence, type of schizophrenia, length of hospitalization, or in production of various Rorschach determinants other than *M*. Extreme introversive and extratensive groups were also employed equated for all factors except the *M*:*Sum C* ratio.

The *S*s were studied in two situations, motor inhibition and spontaneous motor activity. The former called for the writing of a brief phrase as slowly as possible without stopping motion of the pencil. The times taken to write the phrase furnished the data for this task. The second situation involved recordings by an observer of a patient's waiting-room behavior during a 15-minute period. Protocols were rated independently according to a scale of degree of gross motor activity and response to environmental objects. Interrater reliabilities yielded an average *r* of .85.

Results generally tended to support the hypotheses. Thus, the High-*M* group showed significantly longer inhibition time and significantly lower waiting-room activity scores than did the Low-*M* group. The introversive *S*s differed in parallel fashion from the extratensives. When the human movement percepts were rated for relative vigorousness of activity, *S*s who showed predominantly *active M* proved to have longer inhibition time and less waiting-room activity than *S*s producing *static M*. To check the role of color responses High- and Low-*M* groups were subdivided into High- and Low-*Sum C* groups. Although color did not appear materially related to motor inhibition performance, a regular increase in waiting-room activity scores emerged with the increased weight of the color in the *M*:*Sum C* ratio. The results of the investigation corroborate in a limited way some of Rorschach's original formulations and suggest that the interrelations of motor activity, perception of motion, and fantasy behavior offer a fruitful area for experimental research.

Received May 27, 1953.

References

1. Beck, S. J. *Rorschach's test.* Vol. II. *A variety of personality pictures.* New York: Grune & Stratton, 1945.
2. Fenichel, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.
3. Ferenczi, S. Thinking and muscle innervation. In *Further contributions to the theory and techinique of psycho-analysis.* London: Hogarth, 1950.
4. Freud, S. Formulations regarding the two principles in mental functioning. In *Collected papers.* Vol. IV. London: Hogarth Press, 1946. Pp. 13-21.
5. Goldman, A. The effect of motor activity on the autokinetic phenomenon. Paper read at East. Psychol. Ass., Atlantic City, March, 1952.
6. Klopfer, B., & Kelley, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1942.
7. Levi, J., & Kraemer, Doris. Significance of a preponderance of human movement responses in children below age ten. *J. proj. Tech.*, 1952, **16**, 361-365.
8. Lewin, K. *A dynamic theory of personality.* New York: McGraw Hill, 1935.
9. Meltzoff, J., Singer, J. L., & Korchin, S. J. Motor inhibition and Rorschach movement responses: A test of the sensory-tonic theory. *J. Pers.*, in press.
10. Murphy, G. *Personality.* New York: Harper, 1947.
11. Piaget, J. *The language and thought of the child.* New York: Harcourt-Brace, 1932.
12. Piotrowski, Z. A. A Rorschach compendium. In J. A. Brussel, K. S. Hitch, & Z. A. Piotrowski (Eds.), *A Rorschach training manual.* Utica, N. Y.: State Hospitals Press, 1950.
13. Rapaport, D. (Ed.) *Organization and pathology of thought.* New York: Columbia Univer. Press, 1951.
14. Rickers-Ovsiankina, Maria. Studies of the personality structure of schizophrenic individuals. I. *J. gen. Psychol.*, 1937, **16**, 153-178.
15. Rorschach, H. *Psychodiagnostics.* New York: Grune & Stratton, 1942.
16. Schachtel, E. G. Projection and its relation to character attitudes and creativity in the kinesthetic responses. *Psychiatry*, 1950, **13**, 69-100.
17. Siipola, E. The influence of color on reactions to inkblots. *J. Pers.*, 1950, **18**, 358-382.
18. Singer, J. L., Meltzoff, J., & Goldman, G. D. Rorschach movement responses following motor inhibition and hyperactivity. *J. consult. Psychol.*, 1952, **16**, 359-364.
19. Thetford, W. N. Fantasy perceptions in the personality development of normal and deviant children. *Amer. J. Orthopsychiat.*, 1952, **22**, 542-550.
20. Wiezsäcker, V. v. Der Gestaltkreis. *Pflüg. Arch. ges. Physiol.*, 1933, **231**, 630-661.
21. Werner, H. Motion and motion perception: a study in vicarious functioning. *J. Psychol.*, 1945, **19**, 317-327.
22. Werner, H. Perceptual behavior of brain-injured mentally-deficient children. *Genet. Psychol. Monogr.*, 1945, **31**, 51-110.
23. Werner, H., & Wapner, S. Toward a general theory of perception *Psychol. Rev.*, 1952, **59**, 324-338.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

A Comparison of Raven's Progressive Matrices (1938) with a Short Form of the Wechsler-Bellevue[1]

Bert Levine and Ira Iscoe
University of Texas

The Progressive Matrices (1938), and a short form of the Wechsler-Bellevue Form I consisting of the Comprehension, Vocabulary, and Block Design subtests, were given to 60 undergraduates. There were equal numbers of males and females. The average age was 19.1 years. Each *S* was first given the P.M. and then the W-B subtests. Administration and scoring were in accordance with the directions given in the respective manuals. The mean score on the P.M. was 50.8 with an *SD* of 6.6. No significant sex differences were found. A Pearson *r* of .55 was obtained between the total P.M. score and the total weighted score of the three subtests. Individual weighted subtest scores correlated with the P.M. total as follows: Comprehension .21, Vocabulary .48, and Block Design .63. Set E of the P.M., composed of the 12 most difficult items, correlated .52 with the short form of the W-B, and .23 with the Comprehension, .47 with the Vocabulary, and .56 with the Block Design subtest weighted scores. Set E by itself correlated .85 with the P.M. total. The P.M. appears to tap areas of intelligence most closely related to Block Design and not significantly related to Comprehension. As one moves away from cultural loadings on the W-B subtests, the correlation increases markedly. Inasmuch as the P.M. is easily administered, untimed, and demands no speech and only a minimum of motor ability, its possible utility as an intelligence test with people having certain types of sensory and motor handicaps as well as illiterates and foreign born deserves further investigation.

Brief Report
Received December 2, 1953

[1]An extended report of this study may be obtained without charge from Ira Iscoe, Department of Psychology, University of Texas, Austin, Texas, or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. 4159 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Rorschach Performance in Relation to Two Types of Stress[1]

Murray S. Stopol
Teachers College, Columbia University[2]

In the past several years investigators have become more cognizant of individual differences observed in performance under stress. The fact that some persons show a decrement in performance, while others show improvement, and still others have an initial decrement followed by improvement, has led psychologists to attribute these seemingly inconsistent results to such factors as (*a*) the kind of stress situation employed, (*b*) the type of motivation the subject uses, i.e., ego-defensive or need-persistent, and (*c*) the subject's other personality characteristics. The problem of individual reactions to stress is of practical as well as theoretical importance. For example, if we knew more about the nature of these individual differences, it might enable us to predict who would be adversely affected by a stress situation. At the same time, selection procedures might be readily improved in schools, in industry, and especially in military service, where persons are frequently required to perform intellectual and skilled tasks under conditions of distraction or threat.

The most widely used instrument to assess personality variables associated with behavior under stress has been the Rorschach test. Conflicting evidence seems to exist with respect to the relationship between behavior under various types of stress conditions and alleged Rorschach indices of stress tolerance. Williams [19], Waldman [15, 16], Smith and George [13], and Baker and Harris [2] found a definite relationship between Rorschach measures and reactions to stress, while Allee [1], Eriksen, Lazarus, and Strange [9], Deese and Lazarus [6], and Carlson and Lazarus [4] found no relationship. Among the major possibilities that exist in attempting to resolve these differences in results are: (*a*) different aspects of behavior are being measured in these experiments, or (*b*) the validity of the Rorschach indices may be questionable, or both.

Lazarus, Deese, and Osler [11] and Deese and Lazarus [7] have presented a summary of the literature on the effects of psychological stress upon performance, including personality correlates of behavior under stress. Both reviews present numerous studies of performance under stress in which the Rorschach test, as well as other measures of personality, have been utilized. The authors conclude that the majority of these investigations report either negative findings or else present correlations which are statistically significant but are too small to be of any practical value.

Purpose and Hypotheses

The purpose of this study was to investigate the relationship between certain Rorschach measures, presumably concerned with intellectual and emotional control, and performance under two types of stress situations designated as failure stress and task-induced stress (experimental distraction). It was an attempt to examine experimentally the validity of the claim made by Rorschach workers, on the basis of clinical experience and experimentation, that certain Rorschach variables indicate how vulnerable an individual is to stress.

In a previous study, Stopol[3] found that resistance to failure stress was independent of

[1]The present paper is part of a dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Teachers College, Columbia University, 1951. The writer wishes to express his thanks to Drs. Edward Joseph Shoben, Jr. (doctoral committee chairman), Paul E. Eiserer, Lincoln E. Moses, and Laurance F. Shaffer for their invaluable advice and criticism.

[2]Now at Veterans Administration Hospital, Minneapolis, Minnesota and the University of Minnesota.

[3]Stopol, M. S. The consistency of stress tolerance. To be published.

task-induced stress tolerance. Moreover, a cumulative effect of stress was reported from one session to the next, regardless of the type used. Furthermore, no difference was found between the effects of failure stress and experimental distraction if the temporal sequence of presentation of the type of stress was controlled. Consequently, the Rorschach hypotheses under consideration were applied to each type of stress situation separately.

General hypothesis. The Rorschach test is able to differentiate individuals more able to tolerate and individuals less able to tolerate (*a*) experimentally induced failure stress and (*b*) task-induced stress (experimental distraction).

Specific hypotheses.[4] The following specific hypotheses were tested:

1. The greater the amount of *Z*, the greater the tolerance for stress.
2. Individuals with *R* between 25 and 50 will be more able to tolerate stress than individuals outside that range.
3. Individuals with *F%* between 25 and 50 will be more able to tolerate stress than individuals outside that range.
4. Individuals with *F+%* between 69 and 89 will be more able to tolerate stress than individuals outside that range.
5. The higher the *F+%* on the color cards, the greater the tolerance for stress.
6. The greater the preponderance of *FC* over *CF* + *C*, the greater the tolerance for stress.
7. The greater the preponderance of form-dominant responses over form-subordinate responses, the greater the tolerance for stress.
8. The fewer the number of color shock signs, the greater the tolerance for stress.
9. Individuals with *Av. T/1R* between 9 sec. and 16 sec. will be more able to tolerate stress than individuals outside that range.
10. The higher the mean rank based on a combination of low *W%*, high *Dd%*, and the number of form-dominant responses greater than form-subordinate responses, the greater the tolerance for stress.
11. Individuals who show a coarctative or extratensive pattern will perform less well under stress than individuals who show an introversive or ambiequal pattern.
12. The greater the number of Davidson's [5] signs of effective adjustment, the greater the tolerance for stress.

[4]A detailed description of the criteria used in scoring the various Rorschach measures under consideration may be found in the original report [14].

The specific hypotheses used in this study stem from three sources. The variables used in hypotheses 1, 2, and 9 were suggested by Williams [19]. Hypotheses 3, 4, 6, 10, and 11 were largely based on the results of previous investigators [2, 10, 13, 15, 16, 19] who found one or another of the particular measures in question significantly related to performance under various types of stress conditions. Hypotheses 5, 7, 8, and 12, on the other hand, were derived from clinical experience. In general, the Rorschach factors under investigation are supposedly related to intellectual and emotional control, the necessary characteristics presumably involved in the adequate functioning of an individual in a stress situation.

Method

A detailed description of the method is given in a previous paper[5] and the original report [14]. The Rorschach test, administered according to the directions outlined by Beck [3], and a series of five Wechsler-Bellevue Form II Digit-Symbol Test (DST) [17] practice trials were administered to 48 male Columbia University students during the first session of the experiment. The second session, which took place within 24 to 48 hours after the first, consisted of two periods of three phases each. During the first period in Phase *a*, each *S* was given 10 DST trials under optimal conditions. In Phase *b*, either failure stress or experimental distraction was administered during five of the 10 trials. Phase *c* was a poststress phase where no stress was induced during a block of eight trials. The second period began after a 15-minute interval. Each *S* underwent a procedure similar to the first period except that a different form of stress was induced. Two types of stress conditions were utilized: (*a*) failure stress, which consisted of a series of admonishing and reproving comments strongly implying disapproval of performance, and (*b*) task-induced stress or experimental distraction, which involved a series of bells, a buzzer, and a flashing light which were presented during a given DST trial.

The stress score utilized was the degree of change in performance on the DST during the stress trials as compared to prestress level of

[5]See footnote 3.

functioning within each period. This score was calculated from the mean of the last five pre-stress trials minus the mean of the five stress trials, divided by the range of the last five pre-stress trials.

Results

Before reporting the actual results, a word might be in order concerning the statistical techniques used in testing the hypotheses. The method of rank-order correlation was used in testing hypotheses 1, 5, 6, 7, and 10. The chi-square method was applied in the remaining hypotheses. For hypotheses 8 and 12, chi square was used rather than rank-order correlation because of the narrow range of scores as well as the presence of many ties. The cut-off point was made as near the median as possible. A single-tailed test of significance was used in all instances.

In applying the chi-square test to the data in this study, it was noted and anticipated that the expected cell frequencies would be rather small. It was therefore planned beforehand to calculate the p values of chi square by Fisher's exact method if an obtained value of chi square, corrected for continuity, was in the predicted direction and had a p value of .10 or less. This is in line with Edwards' [8] suggestion.

To check on the scoring reliability of the various Rorschach factors used in this study, every third protocol was given to a qualified Rorschach examiner for independent scoring.[6] The total number of responses scored was 677, involving 779 determinants. Both main and additional responses were tabulated. The scoring reliability for all measures was extremely high, no percentage of agreement falling below 90. Thus we may readily consider the scoring of the Rorschach protocols in this study to be highly reliable.

For purposes of analysis, *S*s having task-induced stress initially, followed by failure stress will be referred to as Group A. The *S*s having failure stress initially, followed by task-induced stress will be called Group B. Since Group A was found to be significantly different from Group B as a result of the cumulative effect of stress, regardless of the type used,[7] each group was necessarily treated separately.

For Group A, no significant relationship was found between any of the hypothesized Rorschach variables and resistance to either failure stress or experimental distraction.

For Group B, the only significant relationship found was between the combination of low *W%*, high *Dd%*, and a greater number of form-dominant than form-subordinate responses and the ability to withstand both failure stress and experimental distraction. The rank-order correlation coefficients were .37 and .47 respectively. Both coefficients are significantly different from zero beyond the .05 level using a one-tailed test. This Rorschach measure was taken literally from Waldman [15, 16] who found this combination to differentiate "slow" from "rapid" learners under stress. The score was obtained by taking the mean rank of each person on each factor in the combination. Thus, an individual who was ranked low on two factors and extremely high in another may be above the median when these three factors are averaged. It is, therefore, somewhat misleading, if not hazardous, to generalize about a combination of *low W%, high Dd%,* and a *greater* number of form-dominant over form-subordinate responses. The legitimacy or justification of using such a measure may thus be seriously questioned. In examining the data further, if we classify each person on the basis of being above or below the median value in *W%, Dd%,* and form-dominant greater than form-subordinate responses, in Group B seven out of 24 *S*s showed all three signs in the predicted direction; in Group A only six *S*s presented all three signs in combination. Of the six *S*s in Group A, none was above the median value for task-induced stress tolerance; three cases were above the median value for failure stress tolerance. Of the seven *S*s in Group B, four were above the median value for both failure stress and task-induced stress tolerance. It appears obvious that this hypothesis fails to hold when subjected to a more stringent and meaningful test. Because of the vagueness of the proposed criteria, and since this hypothesis was not tenable for Group

[6] The writer would like to thank Dr. Shabse H. Kurland, who so generously consented to independently score the Rorschach protocols. Prior to giving the protocols to Dr. Kurland, various scoring samples from both Beck [3] and the remaining 32 records were discussed in detail.

[7] See footnote 3.

A, it is highly possible that the significant correlations obtained with Group B may reflect only chance factors. Furthermore, this suggestion is reinforced when we consider that the probability of obtaining two significant correlations at the .05 level by chance when 24 tests of significance are calculated is .34 [18]. Thus we may conclude that the Rorschach test as used in this study is of no value in predicting efficiency of performance under conditions of failure and experimental distraction.

Discussion

Only one Rorschach measure, a combination of low *W%*, high *Dd%*, and a greater number of form-dominant over form-subordinate responses, was found to be associated with the ability to withstand both failure stress and experimental distraction if failure stress was followed by experimental distraction. The difficulty of testing this measure and the vagueness of this combination were noted. A more appropriate statistical analysis found these factors to have no diagnostic utility, since they could have resulted from chance. Therefore, for all practical purposes, it may be said that no Rorschach modality was consistently found to be associated with behavior under either type of stress situation.

Failure to obtain significant positive results with those Rorschach factors allegedly related to stress tolerance is open to several interpretations. One possibility is that these particular Rorschach measures possess no utility as predictors of failure stress or task-induced stress tolerance. The previous investigators who did find significant relationships may simply have been capitalizing on chance factors. It may also be possible that the Rorschach is able to differentiate between persons with respect to tolerance for a given stress, but by measures other than those hypothesized in this investigation. Another possibility may be that these Rorschach signs are not sensitive enough to discriminate between individuals subjected to an experimentally induced stress which is probably not as severe or complex as some extra-laboratory stresses of a longer duration. Still another possibility may be that the results are a function of the *S*s investigated. It should be further noted that the kinds of behavior which these Rorschach factors are attempting to predict may not be sufficiently unitary to be predictable by an instrument such as the Rorschach, and especially by the so-called "sign approach." If we may speculate, the latter possibility appears to be the most feasible of those suggested. It has been clearly demonstrated that a person's tolerance for stress, to a large extent, depends upon both the quality of the particular stressful stimulus condition to which he is exposed[8] and the type of motivation mobilized in a particular type of stress situation [7, 11]. Furthermore, Sherman and Bell [12] suggest that tolerance for a given stress may be a function of many conditions. If this is the case, stress tolerance may be regarded as a complex phenomenon, nonunitary in nature. Moreover, prediction of such behavior seems almost impossible in terms of a one-to-one relationship with other variables, as was the situation in this investigation. Nevertheless, having no logical basis for selecting any of these possibilities as holding true, we must regard these results as indicating that a number of Rorschach signs presumably associated with resistance to stress fail to meet this particular test of experimental validation.

Summary and Conclusions

The purpose of this study was to determine whether certain measures derived from the Rorschach test were able to predict the ability of persons to withstand two distinct types of stress conditions designated as failure stress and task-induced stress (experimental distraction). A series of 12 hypotheses was evolved.

The Rorschach test and a series of Digit-Symbol Test trials were given to a group of 48 *S*s in two sessions. The second session consisted of two periods, each involving a prestress, a stress, and a poststress phase. Twenty-four *S*s received failure stress initially, followed by task-induced stress, while 24 *S*s experienced task-induced stress initially, followed by failure stress.

Except for a combination of low *W%*, high *Dd%*, and a greater number of form-dominant over form-subordinate responses, which was correlated with the ability to withstand both failure stress and experimental distraction if failure stress was followed by experimental distraction, no specific Rorschach measure was

[8] See footnote 3.

found to be significantly related to the ability to resist either type of stress situation. It was shown that the one significant relationship probably reflected chance factors, since the probability of obtaining two statistics significant at the .05 level by chance when 24 hypotheses are tested is .34. It may therefore be concluded that the Rorschach factors employed in this study are of no value in predicting efficiency of behavior under conditions of failure stress or task-induced stress.

Received May 29, 1953.

References

1. Allee, Ruth. Rorschach responses of extreme deviation in an experimental stress condition. Unpublished master's thesis, Univer. of Missouri, 1948. (Cited in F. McKinney, Ruth R. Hines, G. B. Strother, & Ruth Allee. Experimental frustration in a group test situation. *J. abnorm. soc. Psychol.*, 1951, **46**, 316-323.)
2. Baker, L. H., & Harris, Jane S. The validation of Rorschach test results against laboratory behavior. *J. clin. Psychol.*, 1949, **5**, 161-164.
3. Beck, S. J. *Rorschach's test.* Vol. I. *Basic processes.* New York: Grune & Stratton, 1950.
4. Carlson, V. R., & Lazarus, R. S. A repetition of Meyer Williams' study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1953, **17**, 247-253.
5. Davidson, Helen. A measure of adjustment obtained from the Rorschach protocol. *J. proj. Tech.*, 1950, **14**, 31-38.
6. Deese, J., & Lazarus, R. S. The effects of psychological stress upon perceptual-motor performance. USAF *Hum. Resour. Res. Cent., Res. Bull.* 1952, No. 52-19, 1-15.
7. Deese, J., & Lazarus, R. S. The effects of psychological stress upon performance: a theoretical analysis and review of recent developments. Dittoed manuscript, Sept., 1953.
8. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
9. Eriksen, C. W., Lazarus, R. S., & Strange, J. R. Stress and its personality correlates. *J. Pers.*, 1952, **20**, 277-286.
10. Hertzman, M., Orlansky, J., & Seitz, C. P. Personality organization and anoxia tolerance. *Psychosom. Med.*, 1944, **6**, 317-331.
11. Lazarus, R. S., Deese, J., & Osler, Sonia J. The effects of psychological stress upon performance. *Psychol. Bull.*, 1952, **49**, 293-317.
12. Sherman, M., & Bell, Elizabeth. The measurement of frustration; an experiment in group frustration. *Personality*, 1951, **1**, 44-53.
13. Smith, S., & George, C. E. Rorschach factors related to experimental stress. *J. consult. Psychol.*, 1951, **15**, 190-195.
14. Stopol, M. S. An experimental investigation of the consistency of stress tolerance and related Rorschach factors. Unpublished doctor's dissertation, Teachers Coll., Columbia Univer., 1951.
15. Waldman, M. Personality factors in reaction to a stressful psychomotor task. Paper read at East Psychol. Ass., Worcester, April, 1950.
16. Waldman, M. Personality factors and performance under stress in schizophrenics. *Amer. Psychologist*, 1951, **6**, 314. (Abstract)
17. Wechsler, D. *The Wechsler-Bellevue Intelligence Scale. Form II.* New York: Psychological Corp., 1946.
18. Wilkinson, B. A statistical consideration in psychological research. *Psychol. Bull.*, 1951, **48**, 156-158.
19. Williams, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, **11**, 21-29.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Age, Sex, and the Manifest Anxiety Test[1]

A. W. Bendig
University of Pittsburgh

A shortened (50-item) form of Taylor's Manifest Anxiety Scale [1] was given to 497 students enrolled in undergraduate psychology courses. The *S*s also recorded their sex and age on their answer sheets. Raw MAS scores were converted to normalized McCall *T* scores for the subsequent analysis. The *S*s were divided into seven age groups and two sex groups and their scores analyzed by an analysis of variance. Mean MAS scores were also computed for each age and sex group for inspection for trends. Finally, a critical ratio was computed between the mean MAS scores of the Iowa [1] and Pittsburgh samples and an *F* ratio computed from their variances.

Analysis of variance showed no significant differences (at the .05 level) between the sex or the age groups, nor was the sex-age interaction term significant. While there was not a significant difference between the mean MAS scores of men and women subjects, women show slightly higher mean MAS scores in six of the seven age groups represented. A slight curvilinear (parabolic) relationship between MAS and age can be noted for the men with a peak occurring in the 23-25 year age group. This slight curvilinearity is not apparent for women.

A comparison of the mean transformed MAS scores for the Iowa and Pittsburgh samples gave a critical ratio of 2.73 which is significant at the .01 level. The *F* ratio between the variances of the two samples was 1.11 which is not at all significant. However, the difference between the two means was very small.

Apparently, scores on the Manifest Anxiety Scale are not related to sex differences or to age differences over the range represented by our sample of college students. Selecting students on the basis of extreme MAS scores, which is commonly done in comparing the learning rate of "anxious" and "nonanxious" subjects, does not introduce age or sex differences between the subjects comprising the two groups. The significant difference between the mean MAS scores of the Iowa and Pittsburgh samples may be attributed to population differences. Apparently college students demonstrate significantly less manifest anxiety at Pittsburgh than they do at Iowa although the magnitude of this difference is small.

Brief Report
Received December 4, 1953

[1]An extended report of this study may be obtained without charge from A. W. Bendig, Dept. of Psychol., Univer. of Pittsburgh, Pittsburgh 13, Pa., or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. 4158 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Reference

1. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

The Rorschach White Space Response and "Oppositional" Behavior[1]

Albert Bandura
State University of Iowa[2]

The purpose of the present study was to investigate the validity of some of the assumptions underlying the clinical interpretation of the Rorschach white space response (*S*). In general, despite a voluminous clinical and expositional literature, the basic principles of Rorschach interpretation have not been subjected to critical examination, and there have been relatively few attempts at systematic experimental validation [2, 15]. This is as true of the *S* response as it is of the other major scoring categories.

There is fairly extensive agreement among Rorschach workers concerning the psychological values to be assigned to the *S* response. It is assumed to reflect some sort of "oppositional" tendency on the part of the subject. Rorschach tradition holds further that there are several types of overt behavioral expression possible for this inferred oppositional tendency, the type of expression being determined by the basic personality configuration. This configuration is usually defined in terms of the introversion-extroversion continuum. If the *S* responses occur in an *extratensive* personality pattern, the opposition will be directed against the environment, resulting in behavior characterized as "assertive," "stubborn," "negativistic," and "defiant." If they occur in an *introversive* personality pattern, the opposition will be directed inwardly against the self, resulting in behavior characterized as "inadequacy feelings," "self-distrust," and "self-criticism." In an *ambiequal* personality pattern, the *S* responses indicate that the oppositional trend will be expressed in behavior characterized as "ambivalent tendencies" or "indecision."

Validation studies concerned primarily with the *S* response are for the most part lacking. A review of the available evidence on the relationship between the *S* response and oppositional behavior tendencies reveals some contradictory findings. A number of studies have reported a positive relationship between the occurrence of *S* responses and antisocial behavior [3, 12, 16]. However, these studies present some serious methodological defects, and none includes a comparison with an adequate control group. On the other hand, Endacott's [6] study in which delinquent boys were compared with several normal control groups failed to show any significant differences among the groups in the frequency of occurrence of *S* responses. Two studies specifically designed to evaluate the validity of the *S* response [7, 10] have recently been reported. The findings of both were essentially negative. However, the adequacy of some of the criterion measures used in these studies is rather questionable. In summary, the validity of current interpretations of the *S* response is still to be established.

Statement of the Problem

The present investigation is concerned with the validity of the inferences drawn from the *S* response as reflecting an "oppositional" tendency. Teachers' ratings were used as the independent criterion measure against which to evaluate the Rorschach hypotheses. The specific hypotheses subjected to empirical test were:

a. A significant and positive relationship exists between the number of *S* responses and ratings of "negativism" in extratensive subjects.

b. A significant and positive relationship exists between the number of *S* responses and ratings of "inadequacy" in introversive subjects.

[1]This paper is a modification of a portion of a doctoral dissertation submitted to the Department of Psychology at the State University of Iowa. The writer wishes to express grateful acknowledgment to Dr. Arthur L. Benton for his direction of this investigation and to Dr. Harold P. Bechtoldt for many helpful suggestions.

[2]Now at Stanford University.

c. There is no systematic relationship between ratings of "inadequacy" and the number of *S* responses in extratensive subjects and, similarly, there is no systematic relationship between ratings of "negativism" and the number of *S* responses in introversive subjects.

Method

Subjects. A total of 81 University of Iowa High School students, of whom 33 were females and 48 were males, served as subjects (*S*s). They ranged in age from 13 to 18 years, the median age being 15 years.

Procedure. The *S*s were tested individually. The Rorschach administration differed from the conventional technique only with respect to length of exposure time. While in the standard procedure *S* is usually permitted to look at the cards for as long or short a time as he desires, in this experiment each *S* was required to look at each card for a four-minute period. The Rorschach inquiry was conducted in the conventional manner, i.e., immediately after the *S* had given his response to all ten of the cards.

Rorschach scoring procedure. The Rorschach records were scored according to the scoring system of Hertz. In computing the *S* score, primary space responses were counted as 1 and secondary space responses were counted as .5. The color and movement responses were used to determine the experience types which were defined as follows:

Extratensive: $\Sigma C > M$
Ambiequal: $\Sigma C = M$
Introversive: $M > \Sigma C$

The Rorschach experience-type classification was used in checking the hypotheses concerned with the behavioral inferences drawn from *S* responses in the different experience types. Twenty-one *S*s who fell into, or close to, the ambiequal category were excluded from the analysis. Thirty-two of the most extratensive and 28 of the most introversive *S*s were selected for the analysis. One of the introversive *S*s, who gave a five-response protocol and rejected six of the cards, was excluded from the final analysis. The average $\Sigma C{:}M$ value for the extratensive group was 6.2:1.6 while the average $\Sigma C{:}M$ value for the introversive group was 1.9:3.6.

Rating procedures. Five teachers rated the 59 *S*s on four personality variables, viz., negativism, assertiveness, inadequacy feelings, and self-distrust. In terms of the statements in the Rorschach literature, the first two personality characteristics would reflect the behavioral expression of *S* responses in an extratensive pattern, whereas the last two would reflect the behavioral expression of *S* responses in an introversive pattern. Each personality variable was represented on a graphic rating scale by five successive descriptive statements indicating varying degrees of the behavior being rated. Simple descriptions of observable behavior were used, and the maximal rating was placed sometimes to the right and sometimes to the left in order to minimize a possible response tendency on the part of the rater to place *S* in the same category in each of the ratings. The number of *S*s rated by any one teacher varied as follows: 57, 47, 40, 33, and 11. Fifty-four of the 59 *S*s were rated by at least three or four teachers. The remaining five *S*s received two ratings.

The rating procedure consisted of a modified rank-order method. Each rater was presented with a pile of cards containing the names of the *S*s he could rate. The first step was to sort these cards into five consecutive piles, each pile describing a degree of the behavior being rated. Having done this, the teacher then ranked the names within the piles thus providing a final rank order for the group. A rank of 1 was regarded as the highest in the series, i.e., was assigned to the *S* who was considered most negativistic, most assertive, most inadequate, or most unsure of himself.

The assigned ranks were converted into scale values on the base line of the normal curve by Hull's method [9]. The numerical values obtained from the rank data constituted the operational definition of these personality variables. The mean value for each *S* on each of the personality variables was computed to give a final composite score which was used to evaluate the Rorschach hypotheses.

Results

Reliability of the S *score and ratings.* Since the distribution of *S* responses over the first five cards (51%) was found to be comparable to the distribution over the last five cards

(49%), the split-half technique was employed in estimating the reliability of the *S* response. The correlation between the two halves was found to be .72. The estimate of reliability of the whole test, as determined by the use of the Spearman-Brown formula, was .84. The magnitude of this correlation shows the *S* response to be a consistent response tendency.

The reliability of the ratings was estimated by the use of an analysis of variance technique for incomplete sets of ratings, as described by Ebel [4]. The size of the obtained reliability coefficients (negativism, $r = .83$; assertiveness, $r = .79$; inadequacy feelings, $r = .71$; self-distrust, $r = .81$) reflects considerable consistency among the teachers' ratings.

Intercorrelation of the ratings. Of the four personality variables rated, negativism and assertiveness were designed to reflect the behavioral expression of *S* responses in an extratensive pattern while the inadequacy and self-distrust ratings were designed to reflect the behavioral expression of *S* in the introversive pattern, as assumed in the Rorschach hypothesis. The relationship between these personality characteristics, as measured in this study, was estimated by means of product-moment correlations.

A fairly high and positive relationship ($r = + .64$) was obtained between ratings of inadequacy and self-distrust, which suggests that these two personality characteristics are closely related. A less definite relationship is revealed in the correlation between ratings of negativism and assertiveness ($r = + .32$). There were inverse relationships between ratings of assertiveness and self-distrust ($r = - .74$) and between assertiveness and inadequacy feelings ($r = - .45$). The ratings of negativism seem to be relatively independent of inadequacy feelings ($r = + .10$) or self-distrust ($r = + .04$). In general, the size of the intercorrelations among the ratings is contraindicative of the existence of much "halo" effect.

Relationship of S responses to personality ratings. In testing the hypotheses with respect to the behavioral inferences drawn from the *S* response, product-moment correlations were calculated between the *S* scores and the average ratings of negativism, assertiveness, self-distrust, and inadequacy feelings for the two Rorschach experience types. Since the distribution of *S* scores departed from normality in the direction of positive skewness, the $\sqrt{X + .5}$ transformation was employed and the correlations were run on the transformed data. The results of this analysis are shown in Table 1.

The negativism ratings and the *S* scores in the extratensive and introversive groups correlated + .27 and + .44 respectively. Assertiveness was found to be correlated + .26 with *S* scores in the extratensive group and + .13 in the introversive group. No relationship was revealed in the correlation of + .02 between inadequacy feelings and the *S* scores in the extratensive group and + .09 between these measures in the introversive group. An inverse relationship was found between ratings of self-distrust and *S* scores in both the extratensive group (−.23) and the introversive group (− .30).

A Z' transformation was used to test the significance of the differences between the correlation coefficients in the two experience types for the different behavior ratings. In no case did the values of r differ significantly between the two experience types. Since the hy-

Table 1

Table of Correlations between *S* Scores and the Personality Ratings for the Total Sample and for Two Rorschach Experience Types

Ratings	Experience type		Weighted average	*RS*	Partial correlations
	Extratensive (N=32)	Introversive (N=27)	(N=59)	(N=59)	(N=59)
Negativism	+.27	+.44	+.35	+.15	+.34**
Assertiveness	+.26	+.13	+.20	+.15	+.14
Inadequacy	+.02	+.09	+.05	−.12	+.18
Self-distrust	−.23	−.30	−.26	−.31	−.07

**Significant at the .01 level.

pothesis that the two correlation coefficients for each behavior rating were obtained from a common population is tenable, the weighted average of these correlations was obtained by averaging the corresponding Z' values. The correlation coefficients based upon the combined data show the following relationships between the *S* scores and the ratings: negativism, $r = +.35$; assertiveness, $r = +.20$; inadequacy, $r = +.05$; self-distrust, $r = -.26$.

In the group of 59 *S*s, *S* and *R* minus *S* were found to be closely correlated ($r = +.67$). The degree of relationship between *R* minus *S* and the ratings is also shown in Table 1. The highest correlation was obtained between *R* minus *S* and self-distrust ($r = -.31$), indicating that *S*s who are judged to be relatively self-confident tend to give many responses. The partial correlation technique was used to determine the correlation between the *S* score and the ratings with the influence of differences in total responsiveness held constant. The resulting partial correlations showed the following relationships: negativism, $r = +.34$; assertiveness, $r = +.14$; inadequacy, $r = +.18$; self-distrust, $r = -.07$. Since a prediction can be made concerning the probable direction of the correlation, a single-tailed test of significance was used. None of the last three values of r met the criterion of significance at the .05 confidence level. The correlation coefficient of + .34 between *S* and negativism was significant at the .01 confidence level, which indicates a significant but slight relationship between an individual's tendency to give *S* responses to the Rorschach and a tendency toward negativistic behavior, as rated by his teachers.

Discussion

It has been customary in Rorschach work to utilize the experience-type classification as a differential factor in the interpretation of the *S* response. Examination of the correlation coefficients between the ratings and the *S* scores in the two experience types shows no evidence to support the Rorschach hypothesis that the type of behavior to be predicted from the *S* response is dependent on the experience type in which the *S* response occurs. The differences between the correlation coefficients of *S* score and the ratings in the introversive group and the correlations of *S* score and these ratings in the extratensive group were small and statistically insignificant. This finding has obvious implications for clinical diagnostic work. The practice of interpreting the behavioral significance of the *S* score in terms of experience type deserves to be discontinued until cogent evidence for its justification becomes available.

The obtained positive relationship between the frequency of *S* responses and ratings of negativistic behavior offers some support for the Rorschach assumption that the *S* response represents an oppositional behavior tendency. It is obvious, however, from the moderate size of the obtained relationship that the contribution of the *S* response, taken singly, to the prediction of negativistic behavior would be relatively small. Perhaps, if combined with other Rorschach factors which provide additional valid variance, consideration of the *S* response would add to the effectiveness of the Rorschach in predicting negativistic behavior tendencies.

The failure to secure positive findings with respect to the relationships between the *S* response and ratings of assertiveness, self-distrust, and inadequacy feelings could be conservatively interpreted as being merely a failure to disprove the null hypothesis, given the particular conditions of the present experiment. Such a conservative interpretation rests on the assumption that, in contrast to negativism, attitudes of inadequacy and self-distrust are more difficult to evaluate and consequently were not validly measured by the teachers' ratings. However, it may be pointed out that the teachers rated the *S*s in terms of observable behavior, rather than of their own conceptions of the trait being rated, and that the behavioral descriptions were of the type commonly subsumed under these trait classifications. It may be further pointed out that the reliabilities of the teachers' ratings of these traits were as high as the reliability of the ratings of negativism. This, of course, does not assure validity of the ratings, but does at least indicate that there was consistency in the teachers' estimates of the behaviors assumed to be representative of these trait classifications. With these considerations in mind, a stronger interpretation of the findings seems warranted, viz., there is no evidential basis for the practice of inferring

such traits from the occurrence of the *S* response, and the practice should be discontinued until such evidence is forthcoming.

Summary

The present investigation was designed to study the validity of the inferences drawn from the Rorschach white space response as reflecting an oppositional tendency.

Eighty-one University of Iowa High School students were tested individually with the Rorschach test. Ratings on 59 of the *S*s, 27 of whom presented introversive and 32 of whom presented extratensive experience types, were made by five teachers with respect to negativism, assertiveness, inadequacy feelings, and self-distrust. These ratings served as the independent criterion measures against which the Rorschach hypotheses were tested.

The results obtained were as follows:

1. No influence of the experience type on the nature of the relationship between the *S* response and the behavior ratings was observed.

2. Without reference to experience type, the number of *S* responses showed a significantly positive relationship of moderate degree to the ratings of negativism.

3. Without reference to experience type, no significant relationship was obtained between the number of *S* responses and ratings of assertiveness, inadequacy feelings, and self-distrust.

From the findings it was concluded that:

1. The usefulness of the Rorschach experience-type classification in the interpretation of the *S* response remains unsubstantiated. It is felt that the utilization of the experience type as a differential factor in the interpretation of the *S* response should be discontinued until evidence for its justification becomes available.

2. The finding of a significant relationship between ratings of negativistic behavior and the *S* response represents a partial corroboration of the Rorschach hypothesis of the *S* response as reflecting an oppositional tendency.

3. Since there is no evidential basis for inferring inadequacy or self-distrust feelings from the *S* response, it is felt that this practice also should be discontinued until such evidence is forthcoming.

Received April 1, 1953.

References

1. Beck, S. J. *Rorschach's test.* Vol I. *Basic processes.* New York: Grune & Stratton, 1944.
2. Benton, A. L. The experimental validation of the Rorschach test. *Brit. J. med. Psychol.*, 1950, **23**, 45-58.
3. Boss, M. Psychologisch-charakterlogische Untersuchungen bei antisozialen Psychopathen mit Hilfe des Rorschachschen Formdeuteversuches. *Z. ges. Neurol. Psychiat.*, 1931, **133**, 544-575.
4. Ebel, R. L. Estimation of reliability of ratings. *Psychometrika*, 1951, **16**, 507-524.
5. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
6. Endacott, J. L. The results of 100 male delinquents on the Rorschach ink blot test. *J. crim. Psychopath.*, 1941, **3**, 47-50.
7. Fonda, C. P. The nature and meaning of the Rorschach white space response. *J. abnorm. soc. Psychol.*, 1951, **46**, 367-377.
8. Hertz, M. R. *Frequency tables for scoring responses to the Rorschach inkblot test.* (3rd ed.) Cleveland: Department of Psychology, Western Reserve Univer., 1951.
9. Hull, C. L. The computation of Pearson's *r* from ranked data. *J. appl. Psychol.*, 1922, **6**, 385-390.
10. Ingram, W. Prediction of aggression from the Rorschach test. Unpublished doctor's dissertation, Northwestern Univer., 1951.
11. Klopfer, B., & Kelley, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1946.
12. Rabin, A. I. Rorschach test findings in a group of conscientious objectors. *Amer. J. Orthopsychiat.*, 1945, **15**, 514-519.
13. Rapaport, D., Gill, M., & Shaffer, R. *Diagnostic psychological testing.* Vol. II. Chicago: Year Book Publishers, 1946.
14. Rorschach, H. *Psychodiagnostics.* (Trans. by Lemkau and Kronenberg) (2nd Ed.) New York: Grune & Stratton, 1942.
15. Williams, M. The experimental validation of the Rorschach test: I. Experimental correlations. *Amer. J. Orthopsychiat.*, 1952, **22**, 749-754.
16. Zulliger, H. *Jugendliche Diebe im Rorschach Formdeuteversuch. Eine Seelenkundliche und Erzieherische Studie.* Bern: Haupt, 1938.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Acceptance of Self as Measured by Interviews and the Index of Adjustment and Values[1]

Robert E. Bills
University of Kentucky

Several studies have described the reliability and validity of the Index of Adjustment and Values [1, 4], its ability to differentiate personality types [2], and its usefulness as a research tool [3]. In order to relate research findings on the Index to the larger body of research concerning "acceptance of self" as an important personality variable, it should be established that acceptance of self as measured by the Index and by interviews are essentially the same.

Verbatim transcriptions of recorded, 30-minute, open-ended interviews were collected from 13 people. Five of the subjects, all male, were graduate students in educational administration, and eight, six male and two female, were public school administrators.

The subject's total response between two interviewer responses was accepted as the unit to be scored, and the responses were classified according to content and attitude. Statements were classified according to self, others, work, environment in general, and unclassifiable, and type of attitude (positive, negative, and neutral). Ambivalent statements were scored once for plus and once for minus if neither the positive nor the negative aspects appeared predominant, but if either the positive or negative aspects were stronger, then that aspect was given an additional score of plus or minus.

Two judges, working independently of each other, scored all the interviews. Agreement between judges was determined by checking the scoring of each item and noting the number of agreements and disagreements. The percentages of agreement between the judges were 69, 70, 67, 71, 80, 68, 69, 59, 70, 70, 76, 77, and 63, which are satisfactory for the purposes of this study.

When the judges disagreed on the scoring they discussed the difference and arrived at a decision regarding the unit. The revised scoring was used to determine acceptance of self which was calculated from the formula: number of positive attitudes toward self divided by the total of the positive and negative attitudes toward self.

Subjects were ranked according to acceptance of self shown by the interview material and the Index of Adjustment and Values, and the ranks were correlated to give a rho of .84, which is significantly different from zero at less than the .01 level and which permits the conclusion that what a subject says about himself in an interview corresponds highly with the ratings he gives himself on the Index of Adjustment and Values.

Received April 6, 1953.

[1]The data for this study were collected as part of a larger study in the improvement of educational administration sponsored by the W. K. Kellogg Foundation. The University of Kentucky project is directed by Robert L. Hopper, Director of the Bureau of School Services. The writer would like to express his thanks to Henry Oppenheim who has assisted with the project and helped score the interview material.

References

1. Bills, R. E., Vance, E. L., & McLean, O. S. An Index of Adjustment and Values. *J. consult. Psychol.*, 1951, **15**, 257-261.
2. Bills, R. E. Rorschach characteristics of persons scoring high and low in acceptance of self. *J. consult. Psychol.*, 1953, **17**, 36-38.
3. Bills, R. E. A comparison of scores on the index of adjustment and values with behavior in level-of-aspiration tasks. *J. consult. Psychol.*, 1953, **17**, 206–212.
4. Roberts, G. E. A study of the validity of the index of adjustment and values. *J. consult. Psychol.*, 1952, **16**, 302-304.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Prediction of Aggression from the Rorschach

Winifred Ingram
Northwestern University[1]

The prevailing need for experimental validation of the Rorschach prompted the investigation reported here. It is an attempt to appraise the validity of the Rorschach as a measure of aggression, with specific reference to the white-space response, by predicting behavior in an experimental situation. The precision of an experiment with the Rorschach may be considerably lowered when the entire test is the basis of the predictions, since many of the scores which enter into the whole interpretation have not been verified. However, it has been shown that no greater precision is secured by using single or composite scores. Rather, success has been achieved more often where the total test was used [3]. It is recognized that an evaluation of isolated scores or interpretations is, in effect, an evaluation of the whole test. Moreover, it is apparent that it is not possible to forecast what scores can be held accountable for prediction errors when other sources of error which can be controlled have been eliminated. The prediction of actual behavior in concrete situations presents many difficulties; among the major problems are the definition of actual behavior and concrete situation [7], the forecasting of expected behavior, the time allotted to the experiment and its analysis, and the number and kind of procedures used [5]. The experimenter must be particularly aware of the usual cultural modes of expression. Presumably, the socialized individual will be likely to utilize these forms of behavior rather than some more highly individualized acts. The above factors were given major attention in the planning of this experiment and the analysis of the results.

Rorschach states that "*S* answers always indicate some sort of *tendency to opposition*" [6, p. 39]. Beck's interpretation in general adheres to this formulation, but he has added some nuances to the scoring and interpretations. Particularly, the delineation of "unitary reactions" and "true reversals of figure and ground" implies a suggestion of varying personality reactions. Depending upon the configuration of the entire test, Beck has applied the adjectives obstinate, skeptical, ambivalent, and compulsive to those persons responding with a certain amount and kind of *S* response [1, pp. 52-57; 2]. These traits have been reported to accompany or follow frustration by the many investigators who have experimented with the frustration-aggression hypothesis, and it seems well established now that aggression is a major reaction to frustration. If *S* is an indicator of aggressive tendencies, individuals with an average or above average *S* score should respond to a situation which is frustrating to them with more aggressive behavior, quantitatively and qualitatively, than individuals who do not produce *S*. To test this hypothesis it was proposed to predict the aggressive behavior expected from two groups differentiated on the basis of the number of *S* responses produced. The standard for the quantity of *S* was adopted from Beck, who states "In most healthy individuals a minimum number of *S* percepts — from two to four — will be found" [3, p. 47], and Molish, who found the mean number of *S* in normal subjects to be between zero and four.[2]

For the purposes of this study aggression is

[1]Now at the Neuropsychiatric Institute, University of Michigan. The author wishes to express gratitude to Dr. T. W. Richards for his valuable suggestions and criticisms in the preparation of this study and to Drs. T. W. Zeigler, R. D. Rabinovitch, and Mrs. Lois Withey who read this manuscript and suggested many helpful revisions.

[2]Personal communication from Dr. Herman Molish, research psychologist, Psychology Laboratory, Michael Reese Hospital, Chicago, Illinois.

viewed as a dimension with socially acceptable behavior at one extreme and socially unacceptable behavior at the other. From among the adjectives used by Beck and Rorschach to describe the characteristics of the trait projected in *S* and the terms used to specify aggressive behavior by the frustration-aggression theorists, the following were selected as most representative of the trait under consideration: self-assertion, initiative, persistence, resistance, and hostility. The first three were presumed to represent socially acceptable behavior and the last two were considered to belong in the category of socially unacceptable behavior. In addition to these five aspects of aggression, two terms, cooperation and rapport, were selected as representing the opposite of aggression. Operational definitions, an example of which is given below, were derived from observations of behavior in pilot runs of one of the situations, reports of behavior in frustration experiments [11], and familiarity with interviewee behavior. Each set of operations is preceded by dictionary definitions [9].

Assertion

1. the maintaining or defending, as a cause or claim.

2. positive declaration or averment, affirmation; statement asserted; position advanced.

Operation Interview

1. maintains opinion or attitude about himself in the face of opposing opinions.

a. positively by affirming his position, demonstrating or illustrating his position or ability.

b. negatively by denying the accusation.

2. interrupts interviewer to maintain a position about himself.

3. forceful emphasis of the position without accompanying hostility (i.e., a minimal amount of hostility).

Operation Puzzle

1. ignores instructions even when repeated by examiner.

2. repeats same pattern of response even when it does not achieve the goal.

3. asserts his belief in his ability to solve the puzzle.

Other variables were defined in a similar manner.

Method

The present study consists basically of two parts. The first of these is a comparison of two groups of individuals in two life situations in which aggression might be expressed. The second part of the study is concerned with the correlation between aggression as expressed in each of the life situations and in the Rorschach protocol as a whole.

Sample. From the student records in the office of the Dean of Students of a small state college in Ohio 81 male students were selected for testing. Younger students were given priority if their IQ was average or above and there was nothing in their records to indicate emotional maladjustment. It was hoped, by the selection of Negro students, to decrease problems of rapport which might arise because of racial differences between experimenter and subject. Males were expected to be less inhibited in their display of aggression. Younger students with less experience as experimental subjects could be expected to react more naturally. It was felt that a greater degree of cooperation would be insured by eliminating overtly maladjusted students. The IQ range was limited to average or above to decrease the possibility of sparse records.

Certain cultural traits of American Negroes were expected to modify behavior in the experimental situations. An analysis of the subjects' family background and familiarity with their behavior on campus led to the conclusion that respect for authority, expression of feeling by indirection, and a fairly rigid code of etiquette were characteristic for this group. Status might decrease or increase the amount of aggression overtly displayed. The occupational status of parents indicated that the group on the whole was middle class.

Selection of subjects for the final phases of the experiment was on the basis of the number and type of *S* produced. Each test was given a preliminary examination to score *S*, alone, at the end of each day's testing. For the experimental group only subjects who gave two or more reversals, i.e., responses in which the white space was perceived as the figure and the dark or colored portions of the blot as the ground, were selected.

Approximately half of the 38 subjects tested produced some *S*. Eleven of these met the requirements. These students were then paired on the basis of age and Otis IQ with those who produced no *S*. Three pairs were dropped from the group because of leaving school or

erroneous scoring of *S* in their protocols. The final sample consisted of 16 subjects divided equally between the groups. The mean difference in IQ between pairs was three, with a range from nine to five. The mean difference in age was 11 months with a range of 0 to 26 months. The experimental group ranged in age from 18 to 23 (median 19.5) and in Otis IQ from 99 to 117 (median 108.5). The control group ranged in age from 18 to 22 (median 19.5) and in Otis IQ from 99 to 112 (median 108.5). The largest proportion of the total sample was freshmen.

Procedure. The Rorschach test was administered in the usual manner with no exceptions. After all experiments had been completed, the tests were scored by the author according to the Beck method. The scored records and protocols were typed for mailing to the judges.

Two months elapsed between the administration of the Rorschach and the two experimental situations. Two situations were planned which differed in the type of conflict induced; one of them was designed to simulate as closely as possible a life situation. Behavior was recorded by two observers in the first situation (problem solving) and by observers and wire recordings in the second (interview). Unsupervised observers in the first situation were guided in recording behavior by a detailed schedule. In the interview situation the author interviewed while a hidden observer took notes and operated the wire recorder. These two practiced keeping running notes of reactions on two occasions before the subjects were seen.

The problem-solving situation, designed to elicit reaction to intellectual frustration by setting up a goal difficult to achieve, preceded the interview by three to seven days. The Seashore pyramid puzzle [8], selected for this purpose, consists of a three-peg board and six graduated disks with holes which allow them to be slipped easily on the pegs. The goal of the puzzle is to move all six disks from the left-hand peg to the extreme right-hand peg, moving only one disk at a time and never placing a larger disk on a smaller one. Two psychology students served as examiner and observer; the examiner gave the instructions and kept the time and a running record of the subject's behavior, while the observer counted the number of moves and interruptions. A time limit of 30 minutes was set, but the subject was not informed of this limit unless he used all of the time. He was then informed "Your time is up." When the subject failed to observe instructions, the examiner reminded him by repeating pertinent parts of the initial instructions. Questions were answered in the same way. If the subject completed the test in less than 30 minutes, the examiner waited not less than five minutes for the subject to ask to repeat the test. If within this time the subject did not volunteer to repeat, the examiner asked him if he would like to do it again and accepted whatever answer was given. Throughout the situation the examiner exerted a minimum of pressure on the subject.

It was assumed that in the interview situation the interviewer blocked the path to the subject's goal, his vocational choice. Preceding the interview a questionnaire concerning the subject's academic history, vocational choice, extracurricular activities, and family constellation was mailed with a letter to each student stating that the author had completed analysis of the Rorschach tests and wished to interview him. It was requested that the completed questionnaire be returned at the time of the interview [4].

The interview office contained a small closet. The observer sat in this closet, operated the wire recorder, and kept a running record of the subject's behavior. The subject sat next to the interviewer's desk. As the subject was admitted, the observer started the wire recorder so that the sound could become part of the room noises. The interview was usually opened by the interviewer asking for the questionnaire, sometimes by the subject offering it. The interviewer would then read through the questionnaire rapidly and ask the subject, "What difficulty are you having with your classes?" Next the interviewer requested "Tell me more about why you would 'like to be a .....,' " naming the subject's vocational choice as stated in the questionnaire. Following the subject's response, the interviewer pretended to read the subject's Rorschach report and then stated seriously: "After considering your records carefully, your Rorschach, the report of your puzzle performance, and what you have said here, it's difficult to see how you're making

the grade. It looks like a pure waste of time to me." A concealed stop watch was started as this statement was completed, and the time of the subject's first response, verbal or otherwise, was recorded. When the interviewer felt certain that the response to this statement was completed, the time of the last response was recorded. The interview was terminated with an explanation of its purpose; questions were answered and a thorough attempt was made to desensitize the subject before he left.

The three situations provided settings for different roles for the participants, observers, and subjects. Use of the Dean of Student's office and its services gave the author quasi-official status, which was further enhanced by a notice placed on the bulletin board that the author, an employee of the U. S. Public Health Service, was engaged in making a survey of normal young men. The role of an objective psychologist was developed by not using the subjects' names, with the explanation that this enabled the author to remain more objective. Living on campus facilitated friendly relationships. It was assumed that establishing such rapport would reinforce the frustrating effect of the belittling statement introduced in the interview. Although the assistants in the puzzle situation were cast as disinterested observers, there was evidence that a few of the subjects felt some personal involvement with the observers. This is not surprising since the student body is small. Nevertheless it is felt that the subject was pitted against himself rather than against the others in the situation.

Each category of aggression and the assumed opposites were set up as an independent five-point rating scale; all seven scales were printed on a single page. Rorschach, puzzle, and interview protocols were rated by means of the same scales. Judges of the Rorschach records were asked to arrive at an interpretation of the whole record and to record their impression of each subject on the rating scales. There was no training period; instructions and records were mailed. Judges of the Rorschach were selected primarily on the basis of their training in the Beck system.[3] Judges of the other two situations were selected because of their familiarity with rating procedures,[4] a separate group rating each situation. The author acted as a judge for all situations. Definitions and instructions were reviewed for each judge before the ratings were performed. The protocols were identified only by case number, and these were masked before the author acted as a judge.

[3]Drs. Seymour Klebanoff, Herman Molish, Alan Rosenwald, and Meyer Williams acted as Rorschach judges.

Results

Results of the statistical analysis of data will be presented in the following order: (*a*) reliability of ratings of aggressive behavior in the life situations; (*b*) significance of differences between groups high and low in *S*; (*c*) ratings of aggression expressed in the Rorschach performance; (*d*) correlations between ratings of aggression in the life situations and ratings of aggression on the Rorschach records.

Reliability of ratings of aggressive behavior in the life situations. Reliability coefficients were calculated using the total group of 18 subjects.[5] Each coefficient represents intercorrelation between three judges for the puzzle situation and a different group of three judges for the interview situation. For persistence and hostility in the puzzle situation and for initiative and hostility in the interview situation p equaled .01. For initiative, cooperation, and rapport in the puzzle situation and for assertion, persistence, and rapport in the interview situation p was .05. These findings indicate that of the seven behavior variables under con-

Table 1

Reliability of Ratings
($N = 16$)

Behavior category	Situation		
	Puzzle	Interview	Rorschach
Assertion	.15	.58*	.74**
Initiative	.55*	.64**	.70**
Persistence	.59**	.51*	.71**
Resistance	.11	.41	.65**
Hostility	.73**	.71**	.57*
Cooperation	.49*	.27	.72**
Rapport	.56*	.47*	.44

* Significant at .05 level.
** Significant at .01 level.

[4]Miss Alene Gustavson and Dr. Harry Klehr rated interview records. Dr. Ross Morgan and Mr. Dave Murray rated puzzle records.

[5]Calculated by a method suggested by Dr. E. L. Clark.

sideration, four (initiative, persistence, hostility, and rapport) were measured reliably by both the puzzle and interview situations; cooperation was measured reliably only in the puzzle situation and assertion only in the interview situation.

Significance of differences between groups high and low on S. In order to test the significance of differences between groups, ratings for each subject were averaged and the usual *t* test was made. A significant level of differentiation ($p = .05$) between groups was obtained in the interview situation on the following variables: initiative, resistance, hostility, cooperation, and rapport. Differentiation at the same level was obtained in the puzzle situation on assertion and hostility.

Limiting discussion to differences at the .05 level of confidence or better: (*a*) in the interview situation the group high in *S* response on the Rorschach exhibits more initiative, more hostility, but also more cooperation and rapport and less resistance; (*b*) in the puzzle situation the group high in *S* response exhibits less assertion and less hostility.

So far as the interview situation is concerned, the original hypothesis is confirmed, namely that individuals high on *S* response are more aggressive. They are aggressive in being generally more participatory and interactive. However, in the puzzle situation (primarily a situation of intellectual challenge involving social interaction only peripherally) the high *S* individuals are not aggressive. There is a strong suggestion that the two situations differ markedly in the degree to which they elicit aggressive behavior as measured in this study.

Ratings of aggression expressed in the Rorschach performance. Intercorrelations between judges rating the total Rorschach record are shown in Table 1. A probability level of .01 was obtained for the intercorrelations on the variables assertion, initiative, persistence, resistance, and cooperation. The intercorrelations for hostility were significant at the .05 level. These coefficients suggest that the raters were in considerably greater agreement in rating the Rorschach than they were in rating either the interview or the puzzle situations.

Correlation between ratings of aggression in the life situations and ratings of aggression on the Rorschach record. Coefficients of correlation between the ratings of aggression on the total Rorschach and those for each of the life situations are shown in Table 2. The correlation for the variable, persistence, in the puzzle situation is significant at the .05 level of probability. The probability approached the .05 level for the variable, resistance, in the interview situation. A general conclusion is that there is little relationship between the aggression inferred in either situation and that inferred from the total Rorschach. The largest coefficient (.52) suggests a slight correspondence between the puzzle and Rorschach situations for persistence, in the direction which is perhaps expected. In the interview situation, however, the largest coefficient (−.40) suggests that those who seem most resistant on the Rorschach seem least resistant in the interview situation.

Table 2

Correlation between Ratings of the Life Situations and the Rorschach

Category	Situation	
	Puzzle	Interview
Assertion	.34	—.37
Initiative	—.11	.21
Persistence	.52*	.35
Resistance	—.17	—.40
Hostility	—.08	.27
Cooperation	.19	—.01
Rapport	—.30	—.11

* Significant at .05 level.

Summary and Conclusions

In order to demonstrate that the interpretation of the Rorschach white space score, *S*, as aggression is valid, 16 Negro males between the ages of 18 and 24 were subjected to intellectual and interpersonal frustration after having been given the Rorschach test. The resultant observations of behavior in two situations, a problem-solving test and an interview, and the scored Rorschach tests were subjected to a rating procedure on a five-point scale of aggression, composed of five behavior variables, and a nonaggressive scale, composed of two behavior variables.

Of the five aggressive variables the most reliably rated were initiative, persistence, and hostility. Of the two nonaggressive variables

rapport was the more reliably rated.

The major conclusions from the study are:

1. The high *S* group was more aggressive (in a participatory, interactive way) in a frustrating interview situation than the low *S* group.

2. The high *S* group was no more aggressive than the low *S* group in a situation of intellectual challenge in which social interaction was at a minimum.

3. The judges' impressions of aggression as inferred from the total Rorschach record were but slightly related to aggression as measured in the interview and puzzle situations.

Received May 18, 1953.

References

1. Beck, S. J. *Rorschach's test.* Vol. I. *Basic processes.* New York: Grune & Stratton, 1944.
2. Beck, S. J. *Rorschach's test.* Vol. II. *A variety of personality pictures.* New York: Grune & Stratton, 1946.
3. Cronbach, L. J. Statistical methods applied to Rorschach scores. *Psychol. Bull.,* 1949, **46,** 393-429.
4. Murray, H. A. *Explorations in personality.* New York: Oxford Univer. Press, 1938.
5. OSS Assessment Staff. *Assessment of men. Selection of personnel for the Office of Strategic Services.* New York: Rinehart, 1948.
6. Rorschach, H. *Psychodiagnostics.* New York: Grune & Stratton, 1942.
7. Sargent, Helen. Projective methods; their origin, theory and current applications in personality research. *Psychol. Bull.,* 1945, **42,** 257-293.
8. Seashore, R. S. Pyramid puzzle—a useful device for studying thought. *Amer. J. Psychol.,* 1938, **51,** 549-557.
9. *Webster's New International Dictionary of the English Language.* (2nd Ed., Unabridged) Springfield, Mass.: G. & C. Merriam, 1934.
10. Yacorzynski, G. K. An experimental study of the functions of the frontal lobes in man. *Psychosom. Med.,* 1945, **7,** 97-107.
11. Zander, A. F. A study of experimental frustration. *Psychol. Monogr.,* 1944, **56,** No. 3 (Whole No. 256).

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

The Effects of Figure-Ground Contrast Upon Perception as Evaluated by a Modified Rorschach Technique[1]

Arvord W. Belden and E. Earl Baughman
University of Wisconsin

In a recent article [1] results of research designed to clarify the role of stimulus attributes in determining perceptual behavior on the Rorschach test were presented by the second author. Among other things, this research demonstrated a significant increase in perceived movement responses (scored as *M* responses according to Rorschach convention) on a silhouette series of blots in which figure-ground contrast had been maximized. The exact relationship between figure-ground contrast and movement responses, as well as relationships with other dimensions of perception exhibited on the Rorschach test, could not be specified, however. This was due to the fact that there was no systematic variation of contrast in the experiment, but rather only a utilization of stimuli in which contrast was either eliminated or maximized. Examples of these polar extremes can be seen in Fig. 1, where they are indicated as Cards A and G.

The purpose of the present paper is twofold: (*a*) to present data which make possible a systematic analysis of the relationship between figure-ground contrast and perception in a Rorschach type situation, and (*b*) to present an example of how methodology employed with projective techniques may be combined fruitfully with more traditional experimental methods toward the goal of a more complete understanding of perception.

[1]The authors wish to thank Drs. D. W. Fiske, D. A. Grant, G. Calden, and H. A. Page for their criticisms pertaining to this article. Appreciation is also expressed to Drs. F. A. Mote and H. Leibowitz for assistance in constructing materials used in the experiment.

Method

Stimulus materials. A graded series of seven stimulus figures was used in the experiment, the figures differing from one another only in terms of the amount of light reflected. The form of each figure was identical to that of Card II in the Rorschach test. Card II was selected as the basic form for this study because of its demonstrated ability to evoke movement

Fig. 1. Stimulus figures A and G (left to right).

Table 1

Percentage of Difference in Light Reflected between Figure and Ground

Card	Figure-ground difference	Difference between cards
A	0	
		13.1
B	13.1	
		18.6
C	31.7	
		22.0
D	53.7	
		22.5
E	76.2	
		15.0
F	91.2	
		5.7
G	96.9	

responses [1]. The two extremes of the seven stimulus cards are presented in reduced size in Fig. 1. Table 1 indicates the percentage of difference in light value between figure and ground for each card as measured by a Macbeth illuminometer.

Cards B through G were prepared by covering the figure on Card II in the standard Rorschach with a solid coat of black India ink (in order to eliminate color, shading, and internal form effects) and then photographing the card. Actual size prints were then made with different contrast values. As can be seen in Table 1, the difference in contrast between Cards F and G was less than that between Cards B-C, C-D, D-E, and E-F, the latter cards differing from one another in approximately equal step intervals. Card G was included, however, in order to provide a card with maximum figure-ground contrast within the limitations of the photographic process used. Card G will be referred to hereafter as the figure with maximum contrast, and Card B will be referred to as the figure with minimum contrast. Cards F through B were identical to Card G in Fig. 1 except for systematic decrements in figure-ground contrast.

Card A is termed a peripheral line figure and was included in the series in order to provide a stimulus without figure-ground contrast, the white of the figure being the same as that of the ground. It was created by placing a transparent paper over the original Rorschach card and tracing in such a way that the pen was always in contact with both the white ground and either a shaded or colored portion of the card. Photographing then followed the procedure outlined above.

Experimental design and subjects. A total of 140 subjects was used in the experiment, 20 subjects being randomly assigned to each of the seven cards. All subjects were white male students enrolled in an introductory psychology course at the University of Wisconsin. After the composition of the groups had been determined, the groups were compared in terms of age, intelligence, and educational level since these variables are known to affect Rorschach performance. Analysis of variance revealed no significant differences between the groups with respect to these variables.

Testing procedure and treatment of data. All subjects were tested by the first author in order to eliminate examiner differences [2]. The entire testing situation was recorded by wire recorder to insure complete and accurate protocols. Uniform instructions, adapted from Beck [3], were given to each subject as follows:

> You will be given a card with a design on it. Look at the card and tell me what you see or anything that might be represented there. Look at the card as long as you like, only be sure to tell me everything that you see on the card as you look at it. When you are finished with the card lay it down on the desk as a sign that you are through with it.

No encouragement for greater productivity was given after the first response. Upon completion of the free association a standard inquiry was conducted. Each protocol was then summarized by the first author following Beck [3], except for the addition of Klopfer's *FM* score [7]. Seventy-six responses were also scored independently by the second author, perfect scoring agreement being found on 88 per cent of the responses. The scoring categories utilized are listed in Table 2.

The data were analyzed in terms of group comparisons, each group being compared with every other group in respect to each of the scoring categories. Fisher and Yates's [5] exact method was used to test the significance of differences in frequency between groups. The theoretical frequencies used in this procedure were determined by casting the scores of all 140 students into total distributions for each of the scoring categories and choosing a cutting point as close to the median value as possible. It was necessary to use this statistical technique because of two factors: (*a*) the marked skewness of the distributions of most Rorschach scoring categories and (*b*) the small theoretical frequencies in some categories. Such difficulties in the statistical handling of Rorschach data have been reviewed by Cronbach [4].

Two exceptions in the above handling of the data are to be noted, however, involving: (*a*) average time per response and (*b*) average total time per subject. The characteristics of these distributions were such as to permit an analysis of variance technique. This analysis was carried out, in addition to the exact method, for the above two scoring categories.

Results

A table presenting median and mean scores for each group in each scoring category may be secured from the American Documentation Institute.[2] Table 2 indicates significant differences between groups for each scoring category when each group was compared with every other group by the exact method. Figures 2, 3, 4, and 5 present graphs of the relationship between degrees of contrast and 11 of the 16 measures of perceptual behavior used in this study. Five measures (*D, Dd, P, FM,* and *M + FM*) are not graphed since they revealed linear relationships with approximately zero slope comparable to *W* and *M* as seen in Fig. 2. Card A is left discontinuous with the remaining cards in Fig. 2, 3, 4, and 5 for two reasons: (*a*) to emphasize the marked difference between this card and the remaining cards as perceived by the subjects, and (*b*) to make more clear the orderly relationship between degree of contrast and perceptual behavior as long as at least a minimum brightness contrast is maintained. Figures 4 and 5 were constructed on the basis of mean scores for the time categories since the characteristics of these distributions were such as to suggest that the means would be more reliable measures than the medians.

Table 2

Significant Differences between Groups According to Scoring Categories

Scoring category*	Significance level .01	.05
R		A–B
D		A–B, A–G
M	B–A, D–A, F–A	C–A
FM		E–C
Total (–) Form		A–B, A–C
P	B–A, C–A, D–A, F–A	E–A, G–A
S	A–B	A–D, A–F
T/1R		F–C, F–G
T/R	B–E	B–A, B–D
T. Time		B–C

Note.—The first letter in each entry indicates which of the two groups exceeded the other with respect to the scoring category. For example, A exceeded B in productivity (R).

*No comparisons for W, Dd, M + FM, FY, Total (+) Form, and Content Categories reached the .05 level of significance; hence these scoring categories are not listed in the table.

[2]Order Document No. 4076 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photoprints. Make checks payable to Chief, Photoduplication Service, Library of Congress.

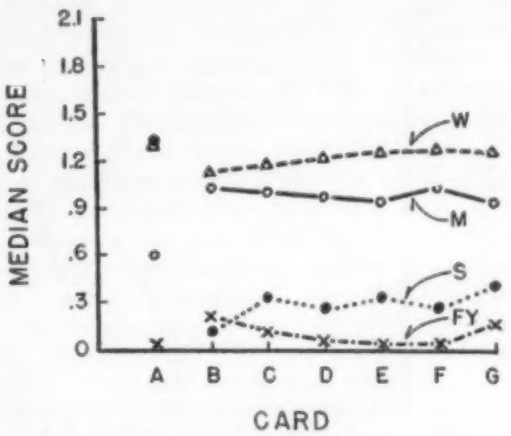

Fig. 2. Median scores for *W, M, S,* and *FY.*

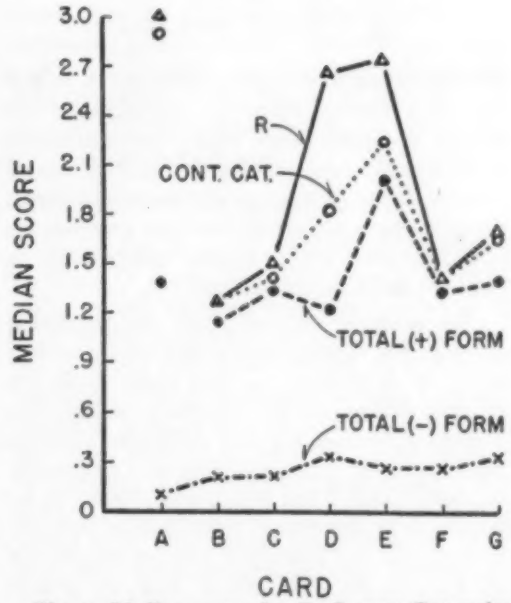

Fig. 3. Median scores for *R,* Content Categories, Total (+) Form, and Total (–) Form.

Nineteen of the 25 entries in Table 2 involve Card A, thus emphasizing the marked changes occurring in perceptual behavior when brightness contrast between figure and ground is eliminated. This is best seen by the fact that the most frequent response given to this card, two humans in some type of activity, dimin-

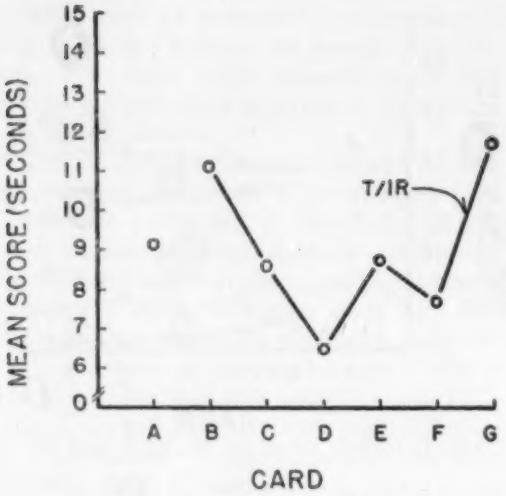

Fig. 4. Mean score for *T/1R*.

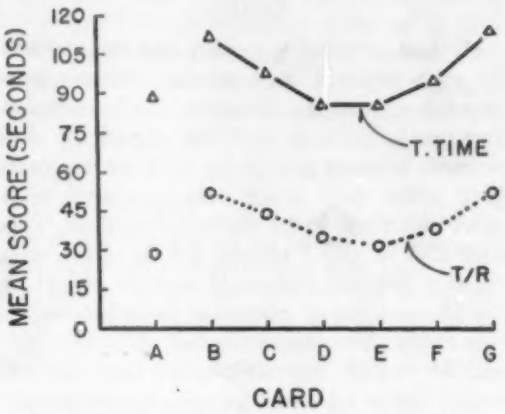

Fig. 5. Mean scores for *T/R* and Total Time.

ishes markedly in frequency on Card A (see *P* and *M* scores in Table 2). In addition, the subject is much more likely to restructure figure-ground relationships on Card A (see *S* score in Table 2) than on the remaining cards. These findings are consistent with previous results obtained when figures comparable to Cards A and G were utilized [1].

Our primary interest from the beginning was with the relationships which would unfold on Cards B through G. On these cards the figure is defined for the subject by means of a brightness contrast, the only difference between the cards being in terms of degree of brightness contrast. Table 2 and Fig. 2 and 3 point to no significant or systematic relationships between contrast and the following scoring categories: *W, D, Dd, M, FM, M + FM, P, S,* and Total (−) Form.[3] Only *W, M,* and Total (−) Form are graphed to illustrate these relationships, thus showing that they may be described essentially by a straight line with zero slope. Figure 3 indicates, however, an increase in three scores (*R,* Content Categories, and Total [+] Form) as contrast increases, until a maximum is reached on Card E, following which there is a decrease in these scores with further increase in contrast. Figures 2, 4, and 5 indicate a reverse type relationship between contrast and four other scores: Total Time, *T/1R, T/R,* and *FY*. These scores reach a minimal point on Card E (except for *T/1R* which reaches a minimum on Card D) and then increase with further increase in contrast.

When *T/R* scores were subjected to an analysis of variance, the resulting *F* was found to be very significant ($F = 5.20$, where $F = 2.95$ is significant at the .01 level). An $F = 2.06$ was obtained when Total Time scores were subjected to a similar analysis, where an $F = 2.17$ would have been significant at the .05 level.

Discussion

Since previous research [1] had demonstrated a marked increase (in comparison with standard Rorschach forms) in movement responses (*M*) on a series of blots with contrast values comparable to those of Card G in the present study, it was anticipated that an increase in *M* responses would be demonstrated corresponding to an increase in contrast values. This relationship was not found, however. A possible explanation for this resides in the uniformally consistent level at which movement responses were evoked by the particular form which was chosen for this study. Thus, the range of behavior with respect to movement responses proved to be too constricted for an

[3] All scores follow accepted Rorschach convention [3] except for Total (+) Form, Total (−) Form, and Content Categories. To obtain Total (+) Form and Total (−) Form scores, all responses were evaluated as to accuracy or inaccuracy of form perceived even though secondary determinants may have been utilized. This was necessitated by the small sample of behavior secured from each individual. Content Categories refers to the number of different types of content perceived by the subject.

adequate test of this hypothesis. The need for further work with diverse forms providing wider variations in perceptual behavior is clearly indicated.

The systematic relationships pictured in Fig. 2, 3, 4, and 5 between contrast and seven of the scores used to evaluate perceptual behavior form the basis of an interesting hypothesis. These curves reveal that more responses (*R*) are given in shorter time (*T/R*, *T/1R*, Total Time) with greater accuracy (Total [+] Form) and with greater variety in content (Content Categories) when figure-ground contrast is intermediate in value rather than minimal or maximal as defined in this study. Moreover, if one accepts the validity of the usual Rorschach interpretation of *FY* responses, the data suggest that feelings of uneasiness or anxiety may diminish at intermediate contrast values. The data are so limited in the *FY* category, however, that only a very tentative suggestion in this direction is justified.

Consideration of the above relationships immediately suggests the idea that perceptual behavior is most efficient when figure-ground contrast is at a level comparable to that provided by Card E (a level at which the figure reflects approximately 25 per cent as much light as is reflected by the ground). As one either increases contrast beyond this point, or decreases it, perceptual behavior tends to become less efficient. This interpretation of the data is consistent with a definition of efficient behavior which stresses maximal output per unit time at a high level of accuracy concomitant with pleasant feeling tone.

The methodology used in this study differs from that of the more usual experimental approach in which both the stimulus conditions and the subject's responses are rather rigidly controlled. In this experiment we have chosen instead to control carefully the stimulus conditions but to make multidimensional measurements of free or unstructured responses by the subjects. This is equivalent to saying to the subject, "If this stimulus condition affects you then you are free to show us how it affects you," rather than, "If this stimulus affects you then it must affect you through the behavior which we have already selected."

An obvious criticism arises immediately, however, in the sense that the many measures used to evaluate the behavior are not measures of independent functions. This problem has been discussed elsewhere by Fiske and Baughman [6]. By ferreting out these interrelations before variations are introduced into the stimulus conditions we should be able to overcome these difficulties. Thus, a previous study [6] has shown that as productivity (*R*) increases there is a correspondingly regular increase in *F+* and *F−* responses. Consequently, the fact that we get such an increase in *F+* responses in the present study but *not* a similar increase in *F−* responses makes more meaningful the interpretation which we have given to the data.

It is possible, of course, that our results are due to the particular form which we chose to work with or to unidentified differences between our groups. Further research will be necessary to determine whether the concept of perceptual efficiency, in the sense that we have proposed it, will hold with other groups of subjects and stimulus forms.

Summary

Seven figures were constructed comparable to Card II in the Rorschach test with respect to form, but differing from one another in degree of figure-ground contrast. Each of these seven cards was administered to a different group of 20 male subjects who were randomly assigned from a total of 140 subjects in elementary psychology courses. Administration and scoring of protocols followed accepted Rorschach procedure. Each group was then compared with every other group in order to determine the relationship between contrast and perceptual behavior, with the following findings:

1. Nineteen out of 25 significant differences in Rorschach responses between cards involved Card A on which there was no brightness contrast between figure and ground.

2. When Cards B through G were considered (on which approximately equal step increases in contrast were present) no relationship between contrast and the following factors was found: *W*, *D*, *Dd*, *M*, *FM*, *M* + *FM*, Total (−) Form, *P*, and *S*.

3. When Cards B through G were considered, *R*, Content Categories, and Total (+) Form were found to increase with increasing

contrast until a maximum was reached on Card E; these scores then decreased with further increase in contrast.

4. When Cards B through G were considered, *T/1R*, *T/R*, Total Time, and *FY* were found to decrease with increasing contrast until a minimum was reached (usually on Card E); these scores then increased with further increases in contrast.

These findings were interpreted in terms of the concept of perceptual efficiency. It was suggested that perception is most efficient at a point of contrast which falls between the minimum and maximum as defined in this study, a point at which the figure reflects approximately 25% of the light which is reflected by the ground.

Received March 30, 1953.

References

1. Baughman, E. E. A comparative analysis of Rorschach forms with altered stimulus characteristics. *J. proj. Tech.*, in press.
2. Baughman, E. E. Rorschach scores as a function of examiner differences. *J. proj. Tech.*, 1951, 15, 243-249.
3. Beck, S. J. *Rorschach's test.* Vol. I. *Basic processes.* New York: Grune & Stratton, 1949.
4. Cronbach, L. J. Statistical methods applied to Rorschach scores: a review. *Psychol. Bull.*, 1949, 46, 393-429.
5. Finney, D. J. The Fisher-Yates test of significance in 2 × 2 contingency tables. *Biometrika*, 1948, 35, 145-156.
6. Fiske, D. W., & Baughman, E. E. Relationships between Rorschach scoring categories and the total number of responses. *J. abnorm. soc. Psychol.*, 1953, 48, 25-32.
7. Klopfer, B., & Kelley, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1946.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Interrelationships Among Several Measures of Anxiety and Hostility [1]

Leonard D. Goodstein
State University of Iowa

The problem of anxiety and its measurement has become increasingly important in contemporary psychology in both clinical and experimental work.

The Taylor Manifest Anxiety Scale. The Taylor Manifest Anxiety Scale (*A* scale) [8, 9, 24, 25] has been widely employed as a measure of anxiety. The *A* scale, initially, consisted of 65 MMPI items that were rated by judges to be indicative of anxiety; those items found to be nondiscriminating, through item analysis, were subsequently eliminated. The present scale consists of 50 such items.

Validity studies of the *A* scale have been of three types. A number of experimental studies have assumed that *A*-scale scores are an index of the *Ss*' state of *reactivity* or *excitability,* which in turn reflects their general level of drive (*D*). Using Hull's [12] theoretical formulation relating response strength to drive, these studies have predicted performance in a variety of laboratory experiments, involving conditioning [1, 10, 16, 21, 22, 24], reaction time [18, 27], verbal learning [15, 17, 23, 26], as well as several more complex behavioral situations [5, 19, 20, 28, 29].

Secondly, several attempts have been made to relate *A*-scale scores to manifestations of clinically diagnosed anxiety states. Gleser and Ulett [6] reported a correlation of .61 between *A*-scale scores and psychiatrists' ratings of anxiety proneness. Lauterbach [14] found a correlation of .31 (significant beyond the 1% level of confidence) between *A*-scale scores and psychologists' ratings of overt anxiety, but the correlation with psychiatrists' ratings was not significantly different from zero.

[1]Based on a paper presented at the annual meetings of the Midwestern Psychological Association, Chicago, Illinois, May 1, 1953.

Thirdly, additional information on the clinical validity of the *A* scale is found in those studies reporting the relationships between *A*-scale scores and other tests purporting to measure anxiety and other personality variables. Lauterbach [14] reported a correlation of .62 between the *A* scale and Welsh's Anxiety Index, an empirically derived MMPI anxiety measure, while Holtzman *et al.* [11] have reported a correlation of .86 with the Winne Scale of Neuroticism, another empirically derived MMPI measure.

The present study is of the third type. It represents an attempt to study the relationship between *A*-scale scores and two other measuring instruments, the Elizur Rorschach Content Test and a new, multiple-choice form of the TAT.

The Elizur Rorschach Content Test. A relatively new technique for scoring Rorschach protocols, based on an analysis of protocol content, has been developed by Elizur [3]. Rorschach responses containing clear-cut evidence of strong anxiety, such as a frightening giant, are given score credit of 2, while responses containing less clear-cut evidence, such as an unpleasant animal, are given score credit of 1. A quantitative score expressing the amount of anxiety is then obtained by totaling the score credits for a particular protocol. A hostility score is obtained in a similar fashion.

Elizur [3] found significantly high correlations between his Rorschach anxiety scores and indicators of anxiety derived from interviews, questionnaires, and rating scales obtained from normal subjects. His anxiety score was also found to differentiate (at the 1% level of confidence) between a group of psychiatrically diagnosed neurotics and a group of normals

who had been matched on the basis of other relevant variables.

The Iowa Multiple-Choice Picture Interpretation Test. Because of the limitations involved in the usual administration and scoring of the TAT, a simpler method of presentation and a more objective method of scoring was devised as part of a research being carried on at Iowa.

Ten TAT cards, 1, 2, 4, 6BM, 7BM, 7GF, 8BM, 13, 14, and 17BM, were selected on the basis of Eron's [4] normative data and personal experience. Four alternative responses were prepared for each of these pictures; each response was composed of a single statement involving one of the following types of themes: anxiety, hostility, achievement, and blandness.

Card 8GF, a young woman sitting with her chin in her hand looking off into space, was used for the purpose of illustration. The four alternatives to this sample picture were:

A. She feels that she cannot go to the party because no one ever asks her to dance. (*Anxiety*)

B. She is annoyed because she has to work while her friends go out. (*Hostility*)

C. She is looking forward to her opening night as the star of a great show. (*Achievement*)

D. She is listening to her favorite radio program. (*Blandness*)

The *S*s are asked to rank order these alternatives according to the interpretation *they* would be most likely to give. Each area or theme is scored as the sum of the assigned ranks. It should be noted, therefore, that the lower the numerical score, the higher the over-all preference expressed for the particular alternative.

While little is known at the present time about the usefulness of these scores, Daily and Hurley [2], in a preliminary study, demonstrated that subjects scoring high on the achievement variable gave significantly more anticipations in a verbal learning situation than those scoring low in achievement. Hurley [13] has essentially confirmed these results.

Hypotheses

The specific hypotheses were:

1. The three measures of anxiety will be positively related.

2. The two measures of hostility will be positively related.

3. The anxiety measures will be negatively related to the hostility measures, as individuals able to express hostility should be less anxious than those unable to do so.

Subjects and Procedures

The three tests, the *A* scale, the group Rorschach (administered according to Harrower's [7] instructions and scored according to Elizur's [3] system), and the Iowa Multiple-Choice Picture Interpretation Test, were administered to 57 college students enrolled in an undergraduate course at the State University of Iowa.

Results

A complete intercorrelation matrix of test scores (including the total number of Ror-

Table 1
Intercorrelation Matrix among All Test Scores
($N = 57$)

Test	Rorschach total *R*	Elizur Anx.	Elizur Host.	Picture Interpretation Test Anx.	Host.	Ach.	Bland
Taylor *A* scale	.18	.38**	—.20	—.17	—.24	.22	.19
Rorschach total R		.25*	—.25*	.14	.03	—.12	—.02
Elizur anxiety			.07	—.06	—.06	—.02	.14
Elizur hostility				—.08	—.02	.02	.02
Picture Interpretation Test†							
Anxiety					*.11*	*—.60*	*—.57*
Hostility						*—.72*	*—.72*
Achievement							*.46*
Blandness							

* Significant at .05 level of confidence.
** Significant at .01 level of confidence.
† The intercorrelations among the Picture Interpretation Test scores are confounded by the rank-order scoring procedures.

schach responses in each protocol) was constructed and is presented in Table 1.

As may be observed in Table 1, a low but significant positive correlation was found between *A*-scale scores and Elizur anxiety scores, indicating some common variance. A low but significant positive correlation was found between total number of Rorschach responses and Elizur anxiety scores while a negative correlation of the same order was found between total *R* and Elizur hostility scores. This would seem to indicate that *S*s who give many responses tend to give responses scored in the anxiety category while those giving few responses tend to give responses scored in the hostility category.

No significant relationships were demonstrated among the *A*-scale scores and the Rorschach scores on the one hand and the Picture Interpretation Test scores on the other hand. The anxiety and hostility scores provided by the Picture Interpretation Test do not seem to tap the same sources of variance as the other measures.

It is difficult to interpret the intercorrelations among the Picture Interpretation Test scores, as these relationships, at least in part, are due to an artifact of the rank-order scoring procedure.

As the *A* scale has usually been found to be nondiscriminating over the middle range of scores, two groups of *S*s (anxious and nonanxious) were selected from the extremes of the *A*-scale distribution. Each group consisted of 10 *S*s whose *A*-scale scores fell, respectively, within the upper and lower 20% of scores for a standardization population of about 2,500 *S*s. The performance of these two groups was then compared on each of the other measures.

The means and standard deviations for the anxious and the nonanxious groups, the resulting *t*'s, and the probability estimates for the *t*'s are presented in Table 2. In the first four cases single-tailed hypotheses were used since the differences were in the predicted direction. The anxious group tended to give more total Rorschach responses, higher Elizur anxiety scores, lower Elizur hostility scores, and give greater preference to the anxiety foil of the Iowa Picture Interpretation Test than the nonanxious group.

These results are in substantial agreement with those reported in Table 2; however, the results based on the extreme groups are statistically more reliable. The relationship between *A*-scale scores and total *R* is essentially a confirmation of the findings of Westrope [29], who reported that anxious *S*s tended to give more Rorschach responses than nonanxious *S*s in individual Rorschach administrations.

Discussion

The present study is part of a continuing series of investigations on manifest anxiety and its behavioral correlates. To date, the Taylor *A* scale has been the principal instrument utilized in these studies. Recently, however, we have become interested in attempting to tap this source of variance by some other technique. The rather restricted range of *A*-scale scores in a college population, the lack of dis-

Table 2
Means on Seven Measures for Anxious and Nonanxious *S*s
($N = 10$ in each group)

Test	Mean		*SD*		*t*	*p*
	Anx.	N-Anx.	Anx.	N-Anx.		
Rorschach						
Total R	27.4	21.6	7.7	5.0	1.9	<.03*
Elizur anxiety	5.2	2.2	2.3	1.3	3.4	<.001*
Elizur hostility	1.2	3.1	1.3	2.9	1.8	<.04*
Picture Interpretation Test						
Anxiety	23.8	26.4	3.5	4.3	1.4	<.10*
Hostility	27.9	31.7	3.8	5.1	1.8	.10
Achievement	22.2	19.3	4.1	3.4	0.9	.40
Blandness	26.1	23.4	4.3	4.6	1.3	.20

* Using single-tailed hypothesis.

crimination over the middle range of *A*-scale scores, and the susceptibility of the *A* scale to faking [9] have, in part, motivated this interest. It was hoped that the Rorschach or a modified form of the TAT would be helpful in this regard. The present study would seem to indicate that the Rorschach might be somewhat more helpful than this modification of the TAT.

One might question whether the relationships among *A*-scale scores and Elizur scores are merely dependent upon the total number of Rorschach responses. Partial correlations, holding total *R* constant, were computed to answer this question. The partial correlation between *A*-scale scores and Elizur anxiety scores, holding total *R* constant, was .36, while the correlation with the Elizur hostility scores, holding total *R* constant, was —.26. These partial correlations are not significantly different from the whole correlations presented in Table 1. The relationship between *A*-scale scores and the Elizur scores does not seem to be merely dependent upon the total number of Rorschach responses, that is, some general productivity factor.

Summary and Conclusions

1. Not all tests thought to measure anxiety measure the same variable. Significant relationships were found between the Taylor Manifest Anxiety Scale and the Elizur Rorschach Content Test of Anxiety. An anxiety score derived from the Iowa Multiple-Choice Picture Interpretation Test, a modified form of the TAT, was not significantly related to the two other measures of anxiety.

2. Not all tests thought to measure hostility measure the same variable. No significant relationship was found between the Elizur Rorschach Content Test of Hostility and the hostility score derived from the Iowa Multiple-Choice Picture Interpretation Test.

3. The relationship between anxiety and hostility depends upon the measures used. A negative relationship was found between the Taylor Manifest Anxiety Scale and the Elizur Rorschach Content Test of Hostility while a positive relationship was found between the *A* scale and the hostility scores derived from the Iowa Multiple-Choice Picture Interpretation Test. The tests of statistical significance, however, do not permit much confidence in these results.

Received May 11, 1953.

References

1. Bitterman, M. E., & Holtzman, W. H. Conditioning and extinction of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 615-623.
2. Daily, J. M., & Hurley, J. R. Need Achievement and behavior in a rote learning situation. Unpublished manuscript, State Univer. of Iowa, 1952.
3. Elizur, A. Content analysis of the Rorschach with regard to anxiety and hostility. *Rorschach Res. Exch.*, 1949, **13**, 247-283.
4. Eron, L. D. A normative study of the Thematic Apperception Test. *Psychol. Monogr.*, 1950, **64**, No. 9 (Whole No. 315).
5. Farber, I. E., & Spence, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120-125.
6. Gleser, Goldine, & Ulett, G. The Saslow Screening Test as a measure of anxiety-proneness. *J. clin. Psychol.*, 1952, **8**, 279-283.
7. Harrower, M. R., & Steiner, M. E. *Large-scale Rorschach techniques.* (2nd Ed.) Springfield, Ill.: Charles C Thomas, 1951.
8. Hedlund, J. L., Farber, I. E., & Bechtoldt, H. P. Normative characteristics of the Manifest Anxiety Scale. *Amer. Psychologist*, 1951, **6**, 504. (Abstract)
9. Heineman, C. E. A forced-choice form of the Taylor Anxiety Scale. *J. consult. Psychol.*, 1953, **17**, 447-454.
10. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94-99.
11. Holtzman, W. H., Calvin, A. D., & Bitterman, M. E. New evidence for the validity of Taylor's Manifest Anxiety Scale. *J. abnorm. soc. Psychol.*, 1952, **47**, 853-855.
12. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
13. Hurley, J. R. Verbal learning as a function of differentially motivating instructions and the achievement motive. Unpublished doctor's dissertation, State Univer. of Iowa, 1953.
14. Lauterbach, C. An empirical study of the Manifest Anxiety Scale and its relationships to other clinical measures of anxiety. Unpublished doctor's dissertation, State Univer. of Iowa, 1952.
15. Lucas, J. D. The interactive effects of anxiety, failure and intraserial duplication. *Amer. J. Psychol.*, 1952, **56**, 54-66.
16. Meyer, D. R., Bahrick, H. P., & Fitts, P. M.

Incentive, anxiety, and the human blink rate. *J. exp. Psychol.*, 1953, **45**, 183-187.
17. Montague, E. K. Role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91-96.
18. Myers, T. Discriminatory reaction performance as a function of test-defined anxiety level and verbal instructions. Unpublished master's thesis, State Univer. of Iowa, 1951.
19. Peck, Ruth. The influence of anxiety factors upon the effectiveness of an experimental "counseling" session. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
20. Rosenbaum, G. Stimulus generalization as a function of clinical and experimentally induced anxiety. Unpublished doctor's dissertation, State Univer. of Iowa, 1950.
21. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp.* Psychol., 1953, **45**, 116-119.
22. Spence, K. W., & Taylor, Janet A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183-188.
23. Taffel, C. Conditioning of verbal behavior in an institutionalized population and its relation to anxiety level. Unpublished doctor's dissertation, Indiana Univer., 1952.
24. Taylor, Janet A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81-92.
25. Taylor, Janet A. A personality test for manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285-290.
26. Taylor, Janet A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61-66.
27. Wenar, C. Reaction time as a function of manifest anxiety and stimulus intensity. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
28. Wesley, Elizabeth L. Perseverative behavior in a concept-formation task as a function of manifest anxiety and rigidity. *J. abnorm. soc. Psychol.*, 1953, **48**, 129-134.
29. Westrope, Martha. Relations among Rorschach indices, manifest anxiety, and performance under stress. *J. abnorm. soc. Psychol.*, 1953, **48**, 515-524.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Mosaic Productions in Cerebro-arteriosclerosis[1]

Brendan A. Maher[2] and Anthony W. Martin[3]
College of Medicine, University of Illinois

Certain patterns of response in Lowenfeld's Mosaic Test have been described as characteristic of organically impaired subjects. Wertham [7], with regard to adults, and Colm [1] on the basis of work with children, report what they claim to be typical signs of organicity as revealed by Mosaic productions. Colm lists four criteria as follows: (*a*) the additive placement of pieces side by side, (*b*) the lining up of pieces of the same shape or color, (*c*) the piling of pieces vertically, and (*d*) the repetition of simple geometric designs.

Wertham describes two general categories of organic impairment to which he applies the terms "cortical" and "subcortical." The Mosaic pattern which is typically produced by the patient with cortically located brain damage is marked, he writes, "by an inability to achieve good configuration, with a dissolution and dismemberment of the Gestalt." The subcortical pattern is characterized by "stone-bound" designs, where at the "expense of an inner plan the patient follows the impetus inherent in the shape and color of the pieces put down so that the whole response becomes reduced to a mechanical or automatic level." In addition he provides one diagrammatic illustration [7, p. 248].

Elsewhere the Mosaic Test has been described as refractory to statistical analysis. A survey of the literature such as that recently completed by Dorken [2] indicates that the statistical testing of diagnostic hypotheses in this test has not yet been undertaken. Normative data have been provided by McCulloch and Girdner [4], Reiman [5], and Stewart and Leland [7] covering various aspects of the finished designs. However, so far as can be ascertained no hypotheses have as yet been clearly formulated and tested. Kerr [3] has reported a validity study based upon the matchings of Mosaics with over-all personality descriptions but gives no indication as to the nature of the interpretive principles used.

The study reported in this paper represents an attempt to define objective scoring criteria for some aspects of the Mosaic Test and to examine their validity. In order to avoid the ambiguities of definition involved in the study of psychiatric clinical groups, it was decided first to investigate the validity of the "organic signs" referred to above. In particular it was decided to examine the pattern characteristics described by Wertham. Organic impairment accompanying cerebro-arteriosclerosis was selected for study as this is regarded by Wertham as falling into both the cortical and subcortical categories.

In pursuit of this aim it was necessary to establish such accurately defined criteria that a trained observer might readily decide whether a given Mosaic production fell into the organic pattern or not. As written, Wertham's definitions suffer from some vagueness. The criteria used in this study were therefore stated more specifically as deductions from Wertham's broader terms and based upon his examples. The most striking aspect of the cortical pattern is the failure to achieve a good configuration. The question of what is "good" or "bad" in configurations is predominantly a matter of judgment on the part of the observer and therefore subject to question. However, it is possible to obtain data as to whether or not a patient has, in his own opinion, successfully completed a configuration. It was felt that the patient's report as to his perceived success or failure to complete a configuration offered the best hope of objectifying this principle. This was so defined under criterion 2 in the five criteria mentioned below.

[1]The writers wish to thank Dr. David Shakow and Dr. Alan K. Rosenwald of the Illinois Neuropsychiatric Institute and Mr. B. Grossman of the Drexel Home for their advice and assistance in this study.

[2]Now at The Ohio State University.

[3]Now at Purdue University.

With regard to the subcortical pattern, it can be seen that Wertham's concept of "stone-bound" or stimulus-bound is essentially descriptive of a species of perseverative response. However, it is possible for the subject to be bound to the characteristics of a stimulus in several ways. He may simply keep on placing pieces of the same shape, or of the same color, or both, in direct contact with each other. Alternatively he may try to produce over-all a Mosaic which has the same characteristics of form or color as its constituent pieces. Thirdly, he may perseverate at the level of repeating the process of putting pieces onto the tray with no discernible pattern and no continuity of structure. Criteria 1, 3, and 4 below were formulated to categorize these three possibilities. Experience with the test also suggested that it would be essential to exclude "representative" Mosaics, even though they might structurally meet the other criteria. By the term "representative" is meant any Mosaic which is intended to depict a concrete object, alphabetical or numerical symbol. Exclusion of this type of production was established as criterion 5. With this general rationale in mind the criteria were defined and used as a basis for the judges' selection as described in the next paragraph.

Method

The writers administered the Mosaic Test to 54 residents of a home for the aged in the city of Chicago. A record was made at the time of administration of all verbal comment, spontaneous or otherwise. The method of administration was the customary one wherein the subject is invited to make anything he wishes with the pieces on the tray. On completion a brief inquiry was conducted to determine what the subject had intended to make and whether he felt that he had been successful in his intention. The age of the group ranged from 64 to 91 years with a mean of 78.6 years. All the subjects were volunteers who responded to a request for individuals to cooperate in a study of aging. For record purposes the completed Mosaics were outlined and colored on the paper base. The records from the total group were then submitted to each of five psychologists on the staff of the Illinois Neuropsychiatric Institute. Each judge was asked to sort the completed productions into two groups on the basis of the five criteria: one group to be those which met one or another of the criteria, and the other to comprise those productions which did not meet any of the criteria. The criteria, as indicated earlier, were:

1. All designs or productions which consist of one or more single lines whether regular or irregular as to shape or color of pieces used.
2. All designs where the subject has failed in completing to his own satisfaction the design which he had stated he intended to make. This judgment involves reference to the verbal record sheet.
3. All designs which are "homogeneous" as to the relation of the shape of the pieces used to the final production; e.g., a square made of squares.
4. All designs in which pieces are simply placed side by side with no discernible over-all pattern, irrespective of shape or color of pieces used.
5. A negative criterion. No successfully completed design which is of a representational character is to be included, even though it may fulfill one of the criteria. Thus a "square made of squares" would not be included if the subject had described it as a "tile floor" and had completed it to his own satisfaction.

Of the 54 productions judged, a total of 27 were selected by the judges as meeting one or another of the criteria. Of this number, 20 were chosen by all five judges. The seven ambiguous cases were discarded from the sample leaving a further 27 cases which had been unanimously judged as meeting none of the criteria. There was therefore complete agreement on 87 per cent of the cases judged.

Following the classification of Mosaics, reference was made to the results of the routine physical examination which had been given to these subjects by the institution physician. These examinations had been made before and independently of this study. On the basis of these records the subjects were classified into those having recognizable signs of cerebral arteriosclerosis with deterioration and those not having such signs. Five of the cases in which the report was equivocal were discarded. Those subjects whose Mosaics had been discarded as ambiguous were also excluded, leaving ultimately a total of 46 subjects, 20 of whom showed signs of cerebral arteriosclerosis. From the remaining 26, a further 20 having good health and most easily matched in mean age with the organic group were selected. The

matching achieved was as follows: The mean age of the organic group was 75.9 years, with *SD* of 7.3 years, and of the control group, 76.5 years, with *SD* of 6.8 years.

The groupings made by the judges were then analyzed with respect to the frequency of distribution of "organic" patterns in both groups. The "organic" Mosaic productions were then further subdivided into those judged as meeting criteria 1, 3, or 4, (the subcortical criteria) and those meeting criterion 2 (the cortical criterion). The results are given in Table 1.

Table 1
Distribution of Mosaic Patterns in a Group of Aged

Group	A Criteria 1, 3, and 4	B Criterion 2	C No Criteria
Organic	10	7	3
Control	2	1	17
χ^2	11.7	10.11	
vs. col C*	$p = .001$	$p = .01$	

* Chi square calculated using Yates's correction.

Discussion

The levels of significance obtained support the inference that there is a meaningful relationship between the presence of cerebral arteriosclerosis, as judged by physical examination, and the production of certain patterns on the Mosaic Test. In general the data support the clinically derived interpretive principles in so far as the criteria used in this study can be regarded as legitimate inferences from Wertham's rather broader definitions. However it should be pointed out that the criteria are not all derived from the same level of description of psychological events. Criterion 2 differs from the rest in that it is not confined to a description of the structural features of the finished Mosaic production but deals rather with a failure in planning as reported by the subject himself. Influences covered by the psychological concepts of level of aspiration, anticipation, and perhaps self-concept may be involved here. The relationship of organic disability to changes in these influences remains undisclosed by a study such as this. Nevertheless the clinician concerned with diagnostic problems rather than theoretical issues may find criteria drawn from several levels of description to be useful aids in his work. It is to this end that this report is primarily directed. The more fundamental implications for theories of perception and projection remain problems for further investigation.

Received May 1, 1953.

References

1. Colm, Hanna. The value of projective techniques in the psychological examination of children: the Mosaic Test in conjunction with the Rorschach and Binet tests. *Rorschach Res. Exch.*, 1948, **12**, 216-233.
2. Dorken, H. The Mosaic Test: review. *J. proj. Tech.*, 1952, **16**, 287-296.
3. Kerr, M. The validity of the Mosaic Test. *Amer. J. Orthopsychiat.*, 1939, **9**, 232-236.
4. McCulloch, T. L., & Girdner, J. B. The use of the Lowenfeld Mosaic Test with mental defectives. *Amer. J. ment. Def.*, 1949, **53**, 486-496.
5. Reiman, M. G. The Mosaic Test: its applicability and validity. *Amer. J. Orthopsychiat.*, 1950, **20**, 600-615.
6. Stewart, U., & Leland, L. American versus English mosaics. *J. proj. Tech.*, 1952, **16**, 246-248.
7. Wertham, F. The Mosaic Test: technique and psychopathological deductions. In L. Abt, & L. Bellak (Eds.), *Projective psychology*. New York: Knopf, 1950. Pp. 230-256.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Sexual Symbolism: A Validity Study

Leon H. Levy
The Ohio State University

Dating from Freud's work on dream analysis [1], and perhaps before, there has been present in psychology and in psychoanalysis, in particular, the working hypothesis that certain environmental objects stand in a symbolic relation to the male and female genitalia.

As originally elaborated by Freud, the symbolic relationship was based on certain structural or functional similarities between the symbol and the genital organ symbolized. Thus, elongated and pointed objects, such as knives, telephone poles, chimneys, and noses were assumed to represent the male genitalia, while round, containing objects, such as boxes, key holes, houses, and ships were considered symbolic of the female genitalia.

Although White [6] has claimed that this hypothesis is no longer in good repute, it would seem instead that it has gained a wide, implicit acceptance and has now become, for many, an integral part of their interpretation of projective materials ranging from figure drawings [5] to the Bender-Gestalt [4].

This is not to say that stereotyped interpretations are necessarily made, or that the productions are considered out of context. In one case the absence of a chimney on a house, for example, may be said to indicate that the individual feels inferior to his father in terms of power and authority while in another this may be taken as a sign of feelings of sexual impotence. In either case, however, the object has been assumed to have a certain fixed symbolic relationship with reality, i.e., chimney = penis.

The purpose of the present study is to investigate this hypothesized relationship. The hypothesis to be tested may be stated as follows: Where abstract objects are categorized by a given means as either male or female, those objects which may be described as elongated, pointed, or angular will be categorized as male; those objects which may be described as round and containing in nature will be categorized as female.

Method

The experimental design was developed with two factors in mind. The first was that the research was being carried on in a public school system where protest and resistance might be encountered should the purpose of the research be known. Secondly, from the standpoint of validity, it was felt that an indirect method of a projective nature would be less subject to any conscious resistances or anxieties which the subjects (*S*s) might have concerning sex. Further, by using an indirect method it was hoped that the hypothesis was being tested at the level at which it is assumed to be operating.

The *S*s consisted of 62 children, 30 male and 32 female, in the fifth grade of a public school in a community of 100,000 population. The *S*s were divided into two groups, A and B, of 35 and 27, respectively. There was an equal proportion of males and females in each group.

The hypothesis was tested by two approaches, each designed to operate at a different level of awareness. In the first approach, *S*s in both groups were told that the task they were about to perform was part of an investigation of immediate memory. The *S*s were presented with a sheet of paper on which ten first names were randomly listed, five of each sex. Below the list of names, ten figures were presented in random order. Five of the figures were elongated or pointed and were therefore assumed to be symbolic of the male genitalia; five were rounded and containing or symbolic of the female genitalia.[1] The experiment was divided into three parts. During the first part

[1]Copies of the figures are available on request from the author.

Table 1

Predicted and Observed Frequency Distributions for Any Given Number of Correct Matches

Number of correct matches	p	Predicted frequency	Observed frequency
0	.001	.2	0
2	.10	6.2	2
4	.39	24.6	29
6	.39	24.6	23
8	.10	6.2	8
10	.001	.2	0

Ss were asked to assign one of the ten names to each of the figures, using each name only once. The sheets were then collected and *Ss* not adhering to the instructions were eliminated from the study.

The second part consisted of paired-associates learning. Ten flashcards each containing one of the ten figures and one of the names were exposed for three seconds each in immediate succession. For Group A, the elongated and pointed figures were presented with male names, while the rounded and containing figures were shown with female names. For Group B, figures symbolic of the male genitalia were presented with female names, and the female figures with male names.

Part three consisted of a retention test in which *Ss* from Groups A and B were asked, on the basis of the flashcards, once more to match symbols with names. Matching was done on blank sheets like those used in part one.

The hypothesis being tested may now be stated in terms of the following prediction: 1. In part one, matching of figures and names so that they are both of the same "sex" would be greater for the entire sample than that expected by chance. 2. On the retention test in part three, Group A would show better retention than would Group B.

The first prediction is based on the assumption that in any forced matching situation, those items perceived as related in any fashion will be matched more often than will items where there is no perceived relation if there are no instructions to the contrary. If the hypothesis being tested is valid, there should be certain regularly perceived relationships between the figures and the names.

In Bartlett's terminology [1], one may assume that if the hypothesis is valid, there exists for the *Ss* a schema whereby there is a definite relationship between type of figure and sexuality. Thus, in part two the learning of Group A may be considered as *pro-schema* learning and that of Group B, *anti-schema* learning. The experimental design as a whole may be viewed as a study in negative and positive transfer effects; hence the second prediction based on the assumption that Group A will be subject to positive transfer effects while Group B will be subject to negative transfer effects.

Table 2

Comparison of Groups on Retention Scores

Group	N	Mean	SD	t
A	35	5.7	1.907	1.384
B	27	5.111	1.929	$>.05$

Thus, while *Ss* may not associate names and figures in part one as predicted owing to certain conscious inhibitions, if this association does exist it should affect the learning of the two groups in part two.

Results

Since the first part of the experiment represents a simple sorting situation, it is possible to test the null hypothesis by first calculating the probability of getting any given number of "correct" sorts, translating this into an expected frequency distribution for the experimental sample, and comparing this by means of chi square with the actual frequency distribution. The results are presented in Table 1.

Chi square for this table is 4.714 which with 7 degrees of freedom is not significant and therefore gives no reason for rejecting the null hypothesis.[2] Hence, the first prediction derived from the experimental hypothesis is not validated.

Table 2 presents the data on scores obtained on the retention test in phase three. The difference between the means of the recall scores for the two groups does not reach statistical significance and hence fails to confirm the second prediction derived from the hypothesis.

[2]In calculating χ^2 the two extreme categories at either end were combined because of the low predicted frequencies [3].

While the mean recall scores were not significantly different for the two groups, there would still be some support for the hypothesis if the types of errors on recall made by the two groups were different, i.e., if the errors of Group B could be shown to be due to intrusions of previously determined associations between type of figure and sex. An intrusion for Group B would be any error where the *S* has matched a male figure with a male name and female figure with female name. In other words, these are errors counter to training. In order to evaluate the significance of such intrusions for Group B, it seems appropriate to compare the percentage of their total errors due to such intrusions with the percentage of errors made by Group A which were counter to training or "intrusions." This is presented in Table 3. It will be noted that there is no difference between these two percentages.

Table 3
Analysis of Errors Made on Retention Test

Group	Total number of errors	Number of errors due to intrusions	% of total errors
A	147	56	38
B	132	50	38

Discussion

The results of this experiment do not lend any support to the hypothesis of sexual symbolism as stated at the outset. While it is true that the results of this investigation should not be generalized much beyond its experimental population, its results do have broader significance in so far as the proponents of universal sexual symbolism do not set any limits on its applicability.

Since man is a symbol-using animal, it is quite reasonable to expect him to develop many symbols related to sexuality. However, the results of this experiment would throw into serious question the universality of his symbolism. Blind analysis of projective materials, which has become the current fashion, would certainly appear to be on tenuous grounds in so far as it relies on hypotheses similar to the one which was tested here.

The face validity and tenaciousness of the hypothesis of universal sexual symbolism may be explained by several factors. The most important perhaps is that the population out of which this hypothesis was derived was a treatment population for the most part. As such, it might be expected, first, that they would be more likely to have sexual problems and to utilize symbolic representation than would the general population, and second, should the therapist have certain preconceived notions concerning symbolism, it is quite likely that these individuals would be more susceptible to suggestion and hence provide the therapist with the validation he seeks.

Further, because of the uncontrolled nature of the observations made by therapists as well as the likelihood of selective perception, it is quite likely that this is another example of the operation of partial reinforcement.

Summary

The present investigation was aimed at testing the hypothesis that certain environmental objects by virtue of their structural characteristics may be taken as being symbolic of the male and female genitalia. The *S*s used were normal children in the fifth grade of a public school.

The hypothesis was tested by means of a simple matching situation and a learning situation, both of which involved male and female first names and abstract figures conforming structurally to either the male or female symbol as described in the literature. In the first part of the experiment, *S*s were asked to match names and figures, and in the second part they had to learn to associate certain names with certain figures. The latter represented essentially a test of positive and negative transfer effects as would have been predicted by the hypothesis.

The results of both parts of the experiment failed to lend support to the hypothesis. The results were interpreted as evidence against the existence of a universal sexual symbolism and an indication of the need for extreme caution in the interpretation of symbolism and in blind analysis.

Received April 27, 1953.

References

1. Bartlett, F. C. *Remembering.* London: Cambridge Univer. Press, 1932.
2. Freud, S. *The basic writings of Sigmund Freud.* New York: Modern Library, 1938.
3. Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1942.
4. Hutt, M. L. *A tentative guide for the administration and interpretation of the Bender-Gestalt Test.* U. S. Army, Adjutant General's School, 1945 (restricted).
5. Vernier, C. M. *Projective test productions. I. Projective drawings.* New York: Grune & Stratton, 1952.
6. White, R. W. *The abnormal personality.* New York: Ronald, 1948.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

The Interpretations of Anxiety by Various Psychotherapeutic Schools[1]

John H. De Michele
University of Illinois

Contemporary need for a briefer and more effective psychotherapy has led to many searching revisions of methodology. This trend can be seen in the works of Alexander and French [1], Horney [11], Mowrer [15], Rogers [18, 19], Thorne [21], and others. Valuable as this research is, the intensive study by each school of its particular techniques is a narrow approach to methodology. Since each school claims excellent results, it is possible that there may be common factors leading to this therapeutic success. An analysis by White [22], for example, found that although there are many psychotherapeutic methods, there are only five processes or techniques common to all schools. In order to test this conception, studies comparing the various psychotherapeutic schools on these common basic techniques are needed and might prove fruitful. An encouraging beginning has been made in the realm of the patient-therapist relationship by Fiedler [7], Heine [10], Miller [14], and Quinn [17]. This patient-therapist relationship would approximate White's first and fourth process, i.e., rapport and transference. Fiedler, for example, found an optimal patient-therapist relationship basic and common to many schools of psychotherapy. Another basic method is interpretation. Although studies comparing the relative effectiveness of the various forms of psychotherapeutic interpretation have been made by Grossman [8] and a few earlier investigators [5, 9, 12], no studies have been made comparing the various psychotherapeutic schools on the technique of interpretation per se. The basic technique of interpretation is essentially White's third process, called the *pointing out of feelings*. This process has two forms: (*a*) the Rogerian *recognition of feeling* and (*b*) *interpretation*, which categorical as they are, White states, "are not sharply different from each other" [22, p. 345]. Now, Bordin [2], Collier [6], and Stone [20] have suggested the very important point that we think in terms of a continuum along any basic method, i.e., that each therapist varies in the amount of development that he allows on the continuum of any basic method. Actually, there are many continua which can be found within any broad basic method. Collier, for example, proposes a continuum of "depth" of interpretation. The idea of continua is valuable, since it can lead to the possibility of measuring variations within each of the broad basic methods. The comparative study of various schools of psychotherapy is valuable in that it can lead to the discovery of continua and the experimental testing of basic psychotherapeutic methods.

In this study, interpretation has been selected for analysis. Interpretation may be defined as the inference by a therapist of the attitudes of a patient from cues supplied by facial expression, posture, language, gesture, dreams, play, etc. The broad objective of the experiment will be to discover continua within the basic method of interpretation; to discover quantitative variations among the different methodological groups of therapists in terms of *what* is selected to substantiate an inference, and *how* they arrive at an interpretation.

The Criterion

In order to carry out the broad objective of this study, the criterion selected for interpretation should be as definitive as possible. An intense degree of anxiety was selected as the criterion, because its somatic manifestations are

[1]This paper is an outgrowth of the thinking of Dr. L. I. O'Kelly, to whom I am grateful for constant stimulation and encouragement.

directly observable physiological reactions. The term *anxiety* is used to represent all of its forms, i.e., anxiety neurosis, prodromal anxiety, normal anxiety, or fear, since the symptoms are the same. While differences between these forms may be inferred or observed situationally, internally, or in terms of feelings, Brown [3, p. 131] and May [13, pp. 190, 203] agree that the directly observable symptoms are the same. A subsidiary objective of the experiment will be to determine which of the most common directly observable symptoms are used to substantiate an inferred intense degree of anxiety.

From the most complete list [4, p. 48] of anxiety symptoms available, a selection of directly observable symptoms was made. To these were added difficulty in sleeping, because its effects can be noticed, and fear, because its intimate relationship should have the highest correlation with anxiety. The symptoms selected were as follows:

1. Jumpiness to slight sounds, lights, etc.
2. Great tension.
3. Agitation to the point of trembling.
4. Great difficulty in breathing.
5. Profuse sweating.
6. Sleeps very poorly.
7. Great fear.
8. Great difficulty in concentration.
9. Great difficulty in speaking.

The Media

Since the symptoms of anxiety are directly observable, and since the inference of anxiety is essentially the judgment of an emotional expression, 10 pictures of people showing various degrees of emotional expression were selected as vehicles which might be conducive to inferring anxiety. The pictures ranged from nondescriptive expressions to intense expressions suggestive of fear, anxiety, or terror. The pictures with the intense emotional expressions of fear were selected on the basis of similarity to the typical ones used in earlier studies concerned with the judgment of emotion [16, pp. 365-369]. Each picture was of one person in a relatively vague background, so that little could be inferred from the situational context.

The Preliminary Experiment

A preliminary experiment was conducted for two purposes: (*a*) to see if judgment of the symptoms of anxiety has a good correlation with a judgment or inference of anxiety, and (*b*) to see if the expected range of emotional expression, subjectively selected by the criteria of facial and postural expressions of emotion, is as diverse as the measured range of the emotional expression suggested by the pictures.

A group of 17 psychologists, composed of clinicians from the Veterans Administration Hospital at Downey, Illinois, and counselors from the Chicago Undergraduate Division of the University of Illinois, was selected for the preliminary experiment.

Each psychologist rated the selected symptoms of anxiety arranged in five-point scales for each of the 10 pictures shown. He then had to rate anxiety and two other clinical syndromes on similar five-point scales. The median rating was found for each of the 10 pictures, and each was correlated by Spearman's rho with the median anxiety ratings for each of the 10 pictures. The correlation coefficients for each symptom with anxiety ranged from .35 to .89. The mean of all of the median ratings was found and correlated .87 with anxiety. We can conclude from the evidence that clinicians who rate these symptoms high tend to rate anxiety high.

The only quantitative indication that the emotional expressions of the persons shown in the pictures are well dispersed is the median score of the amount of anxiety seen by the clinicians in each picture. Pictures 2, 6, and 10 were always judged to have a low degree of anxiety, while picture 1 was always judged to have a high component. Pictures 4 and 9 were judged to have a high degree of anxiety two out of three times, while pictures 3, 5, 7, and 8 varied within the middle area of the scale.

Methodological Types of Therapists

Two distinct methodological schools may be identified as Rogerian psychotherapy and Freudian psychonanalysis. Adherents of these schools should show the greatest methodological consistency. The "eclectic" psychologists and psychiatrists, although each has a distinct training, are less homogeneous. The groups used in this study consisted of 10 Rogerian psychotherapists from the Counseling Center of the University of Chicago, 10 psychoanalysts who were trained at the Chicago Institute of

Psychoanalysis, 6 eclectic psychologists, and 5 eclectic psychiatrists. An effort was made to select the most efficient and experienced therapists; five years of therapeutic experience was generally required.

Procedure

The pictures showed a person expressing some degree of a feeling, just as a person might in a face-to-face therapeutic situation. In section 1 of the directions, the therapist was asked to infer what he thought the person in the picture was feeling. In section 2, the therapist was then asked what in the picture led him to his impression; he was further asked what other signs, cues, or symptoms he would *expect* the person to display. In the case of anxiety this procedure led to a selection of the most diagnostic signs that each therapist uses to infer anxiety. Section 3 of the directions consisted of rating the signs in terms of the magnitude of their contribution to the major impression, but this rating was not too successful. After all of the pictures had been completed, section 4 of the directions was presented; here the therapists had to classify each of the reshown pictures into one of four affective states: (*a*) normal, (*b*) neurotic, (*c*) psychotic, and (*d*) anxiety. In addition, he had to rate each of these states in terms of intensity, from 1 to 4. Section 4 was of value in two ways: first, to check the consistency of the therapist. If he saw anxiety in the initial inference he should also see it in the final classification; and second, to check the relative intensity of the anxiety classification.

The order of presentation was maintained constant, and the therapist had to comply with at least sections 1 and 4 of the directions. Each therapist had to write his own test. The time given to the administrator by the therapist was of major importance in the flexible use of the other two sections of the directions. Under pressure of time, only the material directly relating to anxiety was completed. The therapists were especially urged to do their best when they inferred anxiety.

Qualitative Results and Observations

Most of the therapists were cooperative, interested in the test, and readily made available their facilities. Many psychoanalysts, however, were rather defensive. The writer found many interesting reactions. Some therapists were efficient and quickly understood the directions; others had to have the directions read over and over again. Two therapists apparently experienced anxiety reactions when they could not cope with a picture. These therapists could not go on with the test and appeared very agitated; one perspired freely, while the other had to leave the room for a smoke to settle his nerves. Some therapists thought the test was difficult, others thought it simple. As a whole the therapists disliked the idea of writing the test and would have preferred a recording machine. The protocols of many therapists showed conflicts, and some stated that the test was a "personal" experience for them.

A sample of an excellent handling of the criterion in picture 1 by an eclectic psychologist is shown in the following protocol: The woman is distressed and frightened by something which she sees happening at her right. She is clutching her coat (1), has a puckered brow, and a tense anguished expression about the mouth (1). If one could see this woman through time, i.e., in motion, she might be restless or more or less frozen. One might expect *explosive* behavior, a sudden decision to do some particular thing. She would be inclined to vocalize, breathing would almost certainly be disturbed (4); there might be anomalies of perspiration, e.g., (cold perspiration) (3); there might be a high degree of concentration, preoccupation (2). If someone spoke to her she would either not hear them, perhaps be petulant, or turn eagerly to them. (D-3)

In order to typify the group's approach, the following examples of picture 1 are presented.

A Rogerian therapist's protocol, typical of the group, is as follows: "I think she is in real emotional distress, with fear and pain as her prominent feelings. Cues: the knotted brow (2), the clenched hand (1), muscular strain on the face (3). I would expect weeping, very likely sobbing to follow if she started to explore her feelings." (D-4)

A typical eclectic psychologist's protocol is as follows: "A reaction of panic, consisting of fright, sorrow and anxiety. She is immobilized as far as reacting is concerned (4). Facial expression (2) and position of the hand (1); general impression that she is viewing something in the distance (1). She feels numb and cold and at the same time agitated (1). Her face is hot and she is perspiring. (1) Her mouth is dry and her heart is pounding. (1) Tears are about to come." (D-4)

A typical analyst's protocol is as follows: "Apparently a frightened woman — looks unhappy — and some pain I feel from internal sources; (4)

the painful expression of the eyes, (3) hand clutched to the coat, (1) mouth open. She seems to be looking at something which gives this reaction of fright — she might next turn away. However I've seen psychotic patients on walks react like this from no external stimuli. Accounting for my first impression above. Also there is a pseudo-quality or an overreaction of pain which vanished quickly and the patients appeared more relaxed following it. She may cry from this point on; (1) lower her head, (2) turn her head, (3) show beads of perspiration." (D-4)

Quantitative Results

Inference from subjective feelings. This factor appears to lie on a continuum which has crude quantitative verification. When the test was being administered to analysts, the writer became aware of an involved "soul searching" to find cues to substantiate their inferences. On the other hand, the Rogerians and eclectic psychiatrists overtly seemed little involved. Some therapists felt so strongly about this internal source of cues that they had to write them down. When a therapist is so aware of his feelings as cues, we could expect a different kind of interpretation than from a therapist who objectifies the problem. The protocols show four analysts, two eclectic psychologists, one Rogerian, and no eclectic psychiatrists who used these internal sources as cues. Crude as this evidence may be, it certainly should be investigated further.

Inference of pathology. The results from the classification of affective states under section 4 of the directions was arbitrary and in many cases unsatisfactory; yet a scrutiny of the results showed a wide difference in classification by the different groups. A dichotomy of the classifications was made between the normal and the three abnormal affective states. The analysts inferred the least amount of normality in the set of pictures, and conversely the greatest amount of pathology. The Rogerians and the eclectic psychiatrists inferred the least amount of pathology, while the eclectic psychologists lie in between the two poles. The significance of differences between the mean amount of pathology inferred by the psychotherapeutic groups was found with the Mann-Whitney *U* Test. The results can be seen in Table 1.

Table 1

Significance of Differences between Psychotherapeutic Groups on the Inference of Pathology (Using the Mann-Whitney *U* Test)

Groups	*N*	*U*, *U*–*z* score	*p*
1. Rogerians and analysts	19	(*U*–*z*) 2.82	.005
2. Eclectic psychiatrists and analysts	15	(*U*–*z*) 2.82	.005
3. Eclectic psychologists and Rogerians	15	(*U*–*z*) 1.97	.05
4. Eclectic psychologists and eclectic psychiatrists	11	(*U*) 5	.08
5. Eclectic psychologists and analysts	16	(*U*–*z*) 1.29	NS
6. Eclectic psychiatrists and Rogerians	14	(*U*–*z*) .60	NS

The inference-cue ratio. The inference-cue ratio is the number of inferences made as compared to the number of cues used to support these inferences. For example, if a therapist makes four inferences and uses only two cues to support his inferences, then his inference-cue ratio would be four divided by two, or two. Two inference-cue ratios were found. The first ratio was the number of inferences made under section 1 of the directions to the number of cues selected from the picture. For the second ratio, the expected or inferred cues, signs, or symptoms were treated as inferences; thus, this ratio was the number of inferred cues to the number of cues selected or observed in the picture. The inference-cue ratios were all made on picture 1, which was completed by all therapists, and upon which the directions were best controlled. Table 2 shows the in-

Table 2

The Inference–Cue Ratios of the Psychotherapeutic Groups

Therapists	Rogerians		Eclectic psychiatrists		Eclectic psychologists		Analysts	
	I/C*	IC/C*	I/C	IC/C	I/C	IC/C	I/C	IC/C
1.	0.5	0.0	0.8	1.8	2.5	4.0	1.8	0.0
2.	1.0	0.8	1.5	1.1	1.0	0.8	2.0	3.0
3.	1.2	0.0	1.0	1.7	1.1	1.0	0.5	1.5
4.	0.8	0.0	1.1	0.0	1.0	4.0	1.0	2.3
5.	0.6	0.6	1.0	2.0	0.8	0.0	1.3	2.0
6.	1.0	0.6			0.7	2.0	2.0	2.0
7.	1.0	0.0					1.5	1.1
8.	1.0	0.0					2.7	1.3
9.	2.3	1.6					4.0	3.0

* I — inferences, IC — inferred cues, C — cues.

Table 3
Significance of Differences between Psychotherapeutic Groups on the Inference-Cue Ratios (Using the Mann–Whitney *U* Test)

Groups	*N*	*U*, *U*–*z* score	*p*
1. Rogerians and analysts	18		
I/C ratio*		(*U*–*z*) 1.94	.05
IC/C ratio*		(*U*–*z*) 2.74	.006
2. Eclectic psychiatrists and analysts	14		
I/C ratio		(*U*–*z*) 1.58	NS
IC/C ratio		(*U*–*z*) 1.03	NS
3. Eclectic psychiatrists and Rogerians	14		
I/C ratio		(*U*–*z*) .66	NS
IC/C ratio		(*U*–*z*) 2.05	.04
4. Eclectic psychologists and Rogerians	15		
I/C ratio		(*U*–*z*) .35	NS
IC/C ratio		(*U*–*z*) 2.13	.03
5. Eclectic psychologists and eclectic psychiatrists	11		
I/C ratio		(*U*) 13	.08
IC/C ratio		(*U*) 17	.08
6. Eclectic psychologists and analysts	15		
I/C ratio		(*U*–*z*) 1.53	NS
IC/C ratio		(*U*–*z*) .17	NS

* I — inferences, IC — inferred cues, C — cues.

ference-cue ratios of the various psychotherapeutic groups.

From the inference-cue ratios of Table 2, we can observe a progression from the Rogerians, who clearly show an extreme reluctance to infer or interpret at all, to the analysts, who interpret most from the cues they selected from the picture. This progression is an indication of the "depth" of interpretation continuum. The significance of difference of the two inference-cue ratios was found between the various psychotherapeutic groups; the results can be seen in Table 3.

Nosological orientation. For each group of therapists, the number of consistent inferences of anxiety was found, i.e., where anxiety or fear was inferred in section 1 of the directions and confirmed in section 4 with a rating of D-3 or D-4. Ten analysts gave 21 consistent inferences of anxiety throughout the 10 pictures, while nine Rogerians at the other end gave 8 consistent inferences of anxiety. The ratio of the number of times each symptom is used compared to the number of consistent anxiety inferences made by each group may be expressed as a percentage. This percentage shows how often a particular symptom is used to support a consistent inference of anxiety. In addition, Table 4 shows which symptoms are used most often by the various groups to substantiate their inferences of anxiety.

Table 4
Percentage of the Symptoms of Anxiety Used to Support a Consistent Inference of Anxiety

Symptoms	Analysts	Eclectic psychiatrists	Eclectic psychologists	Rogerians
1. Perspiration	57.1	71.4	33.3	0.0
2. Breathing difficulty	28.5	57.1	25	0.0
3. Immobility	14.3	14.3	33.3	0.0
4. Vasomotor difficulty	14.3	28.5	8.3	0.0
5. Tension	9.5	85.5	33.3	0.0
6. Pupils dilated	9.5	28.5	0.0	0.0
7. Confusion	4.8	14.3	8.3	0.0
8. Exhaustion	14.3	14.3	16.6	0.0
9. Rectal urgency	4.7	14.3	0.0	0.0
10. Pilomotor reaction	0.0	14.3	0.0	0.0
11. Speech difficulty	0.0	14.3	8.3	0.0
12. Vomiting	4.7	0.0	0.0	0.0
13. Nervous movements	9.5	14.3	0.0	0.0
14. Collapse	9.5	14.3	0.0	0.0
15. Flight	19.0	28.5	16.6	12.5
16. Cryout or tears	28.5	57.1	25.0	25.0
17. Gross body movements	47.6	0.0	0.0	62.5

The eclectic psychiatrist, whose training is heavily weighted with nosology, would probably substantiate an inference of anxiety with cues which are more nosologically oriented than would the Rogerian, who at the other extreme has little interest in nosology. A continuum in terms of nosological orientation has some experimental confirmation. There is a progression of the number of nosological symptoms used, extending from the analysts and eclectic psychiatrists at one end to the Rogerians at the other. None of the Rogerians inferred any of the nosological symptoms of anxiety, but instead they used gross bodily movements descriptive of the startle pattern. The inferred or expected cues used were crying, screaming, expressing concern, and movements away or toward the anxiety-producing object.

Summary and Conclusions

The aim of this research was an attempt to compare the various psychotherapeutic schools

on the basic method of interpretation. Therapists were to interpret anxiety from pictures showing various degrees of emotional expression. The results were analyzed qualitatively and quantitatively. Four continua in the basic method of interpretation were postulated from the results of the experiment. These continua are tentative and need more refined experimental substantiation. The four continua are:

1. "Depth" of interpretation
2. Nosological orientation
3. Inference of pathology
4. Inference from subjective feelings

Received June 8, 1953.

References

1. Alexander, F., & French, T. M. *Psychoanalytic therapy.* New York: Ronald, 1946.
2. Bordin, E. S. Dimensions in the counseling process. *J. clin. Psychol.*, 1948, **4**, 240-244.
3. Brown, J. F. *The psychodynamics of abnormal behavior.* New York: McGraw-Hill, 1940.
4. Cattell, R. B. *Description and measurement of personality.* Yonkers, N. Y.: World Book Co., 1946.
5. Closson, F. H., & Hildreth, H. M. Experiment in psychotherapy during selection examination. *Nav. med. Bull.*, 1944, **43**, 39-43.
6. Collier, R. M. A basis for integration rather than fragmentation in psychotherapy. *J. consult. Psychol.*, 1950, **14**, 199-205.
7. Fiedler, F. E. The concept of the ideal therapeutic relationship. *J. consult. Psychol.*, 1950, **14**, 239-245.
8. Grossman, D. The experimental investigation of a psychotherapeutic method. *J. consult. Psychol.*, 1952, **16**, 325-331.
9. Haggard, E. A., & Murray, H. A. The relative effectiveness of three "therapy" procedures on the reduction of experimentally induced anxiety. *Psychol. Bull.*, 1942, **39**, 441. (Abstract)
10. Heine, R. W. An investigation of the relationship of change in personality from psychotherapy as reported by patients and the factors seen by the patients as producing the change. Unpublished doctor's dissertation, Univer. of Chicago, 1950.
11. Horney, Karen. *New ways in psychoanalysis.* New York: Norton, 1939.
12. Keet, C. D. Two verbal techniques in a miniature counseling situation. *Psychol. Monogr.*, 1948, **62**, No. 7 (Whole No. 294).
13. May, R. *The meaning of anxiety.* New York: Ronald, 1950.
14. Miller, Helen E. "Acceptance" and related attitudes as demonstrated in psychotherapeutic interviews. *J. clin. Psychol.*, 1949, **5**, 83-87.
15. Mowrer, O. H. *Learning theory and personality dynamics.* New York: Ronald, 1950.
16. Munn, N. L. *Psychology.* Boston: Houghton Mifflin, 1951.
17. Quinn, R. D. Psychotherapists' expressions as an index to the quality of early therapeutic relationships established by representatives of nondirective, Adlerian, and psychoanalytic schools. Unpublished doctor's dissertation, Univer. of Chicago, 1950.
18. Rogers, C. R. *Counseling and psychotherapy.* Boston: Houghton Mifflin, 1942.
19. Rogers, C. R. *Client-centered therapy.* Boston: Houghton Mifflin, 1951.
20. Stone, D. R. Logical analysis of the directive, non-directive counseling continuum. *Occupations*, 1950, **28**, 295-298.
21. Thorne, F. C. *Principles of personality counseling; an eclectic viewpoint.* Brandon, Vt.: Journal of Clinical Psychology, 1950.
22. White, R. W. *The abnormal personality.* New York: Ronald, 1948.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

The Relationships among Projection, Empathy, Reality, and Adjustment, Operationally Defined

Ralph D. Norman and Patricia Ainsworth
The University of New Mexico

The hypotheses of this study are two: (*a*) when operational definitions of projection, empathy, and reality are given, projection will be negatively correlated with both reality and empathy, whereas empathy and reality will be positively correlated with each other; and (*b*) reality and empathy will correlate more closely with adjustment than will projection.

Regarding the first hypothesis, Allport [2], Healy, Bronner, and Bowers [12], and Sears [15] state that projection can occur provided the individual lacks insight into himself. But self-insight requires a degree of realistic perception of oneself. Steinmetz [16] expands the concept of insight to mean the ability to perceive the world as others do. In either case, it may be stated, as does Alexander [1], that projection involves a lack of reality testing. He says, using projection of hostility as an example, "If this hostility becomes conscious it would cause guilt and anxiety, and to avoid this the whole situation may be distorted by saying, 'I do not want to attack him; he wants to attack me.' Such distortion of reality becomes possible only when the ego renounces its contact with reality under the pressure of emotional needs" [1, p. 114]. Thus, it may be postulated that a negative relationship exists between projection and realistic perception or reality.

That projection is also negatively related to empathy and that empathy is positively related to reality may be inferred from logic and previous research. Certainly, empathy, or the ability to put oneself in another's place, requires realism in seeing others. Dymond [6] has shown that insight and empathy are related to each other, and Norman [13] has demonstrated that there are positive relationships among variables he terms insight, self-other identity, and realistic perception of others. As part of the first hypothesis, then, it is the intention of this study to confirm these findings by Dymond and by Norman, using an entirely different technique. It may also be logically inferred that if reality and projection are negatively related, and if reality and empathy are positively related, then empathy and projection should be negatively related.

Regarding the second hypothesis, it may be predicted that adjustment should be more closely tied to empathy and reality than to projection. Projection is a defense mechanism of the ego and inherently implies maladjustment and insecurity, according to Bellak [3]. It is also prominent in paranoid maladjustment, as is well known. Moreover, Dymond [8] found that individuals who possess empathic ability are better adjusted and more secure than those who do not. Norman [13], using sociometric data of acceptance, found positive relationships between acceptance and self-other identity, insight into self, and realistic perception of others.

Subjects and Procedure

The Guilford-Martin Inventory of Factors GAMIN [11] was given to 74 male college students, most of whom were freshmen. This inventory was chosen because it is easy to administer, was standardized on students, and has high reliability. Particularly, the fact that the test has been factor analyzed is especially advantageous to this study because of interest in the operation of the various experimental variables (empathy, projection, and reality) with traits of high factorial validity. Guilford and Martin [11] state that the intercorrelations between the traits are generally low enough to assure useful differences in measurement between the trait scores. By giving the

test in two forms, scores for our three experimental variables were obtained. The first form was administered as usual except that the "?" alternative in answers was omitted. The subjects thus had to answer either "Yes" or "No." Two weeks later, a second test was given to the same subjects, and, instead of answering for themselves, subjects were required to answer the questions as they thought most others of their own age and sex would answer them. For example, Question 8 on the original, "Do you feel that people almost always treat you right?" was changed to read, "Do you think that most people of your own age and sex feel that people almost always treat them right?" It was felt that there would be more uncertainty in answering questions for others and therefore that too many question marks would be used. The test would then be too difficult to deal with for our purposes, and it was for this reason that "?" was omitted from both forms.

Projection. From statements in the literature [1, 2] projection may be defined to include the following aspects: (*a*) it is defensive; (*b*) it is unconscious; (*c*) it involves undesirable characteristics which threaten the ego; (*d*) it occurs when the ego cannot repress the characteristics; and (*e*) it is a self-deception and therefore a reality distortion. An operational approach was therefore formulated to encompass this definition, as follows. The subject, in answering the first form (about himself) denies that he possesses a certain trait. On the second form, however, he says that others do possess this trait. But it is entirely possible that others *do* possess it, and that he is realistic about others and is not projecting. He *is* projecting if the majority (51% or more) of the group, in taking the test for themselves the first time, say that they do *not* have the trait in question. In other words, if the subject says, "I don't have this, but others do," when actually the majority of others say they do not, then projection is occurring. This procedure was done only for those questions which involved undesirable traits since by definition projection involves undesirable characteristics which the subject cannot repress and against which he unconsciously defends himself. He is being unrealistic in attributing the trait to others when they say they do not possess the trait. This approach is supported by Remmers [14] who, in defining projection, says that if operationally there is a significant difference between the actual measurement of the characteristic and the estimate given by the subject then projection is operative.

Each subject was given a point for projection each time he answered a question according to the above formula. Total scores on projection were obtained for only three of the GAMIN factors—A, I, and N. Trait A is ascendancy in social situations as opposed to submissiveness; I is lack of inferiority feelings or self-confidence; and N is lack of nervous tenseness and irritability. Guilford and Martin state that high scores on G, A, I, and N mean generally unfavorable characteristics and low scores mean favorable characteristics [11]. However, G (general pressure for activity) may possibly be equivocal for this study, since it measures either a tendency to engage in vigorous overt action or a disinclination for motor activity. Likewise, trait M (Masculinity) may be equivocal since a high degree of femininity is to be expected among college males [5]. Therefore, G and M were not used, since they are not as ego-involved as A, I, and N.

Empathy. Dymond [6, 7] defines this as an imaginative transposing of oneself into the thinking, feeling, and acting of another, and with Cottrell [4] she says it means to take another's role, place oneself in his shoes, and perceive the situation from his perspective. Grossman [10] refers to it as perceiving correctly the world from the other person's frame of reference. This perhaps involves projection in that one projects oneself into the thinking, feeling, and acting of another. The difference between this type of transposition and true projection lies in the realistic nature of empathy. Remmers [14], as noted above, has said that operationally if there is a significant difference between the actual measurement of the characteristic and the estimate given by the subject, then projection is occurring. However, he says, if the measurement of the characteristic and the estimate given by the subject are the same, then it is empathy. Thus it can be seen that empathy requires a realistic estimate of others whereas projection does not. The difference then between the projection which occurs in empathy and projection as defined for this study

lies in the fact that in empathy the projection is one which is a true estimate. This reasoning is further justified by the fact that Sears [15] found that those who project lack insight, and Dymond [6] says that those who are empathic have insight, for to possess insight implies a realistic view of oneself and others.

Thus the operational definition for empathy used here is as follows. An individual empathizes when he says others possess a trait in question and others actually do have it when speaking of themselves. In other words, the subject in taking the second form of the test says that others possess a trait; if the majority (51% or more) of the group, in taking the test for themselves (first form), agree with him, then it may be said that empathy is occurring. One point for empathy was given each time the subject answered a question according to this formula. There is then, as Remmers [14] says, no difference between the estimate of the measurement of others and the actual measurement of others.

Reality. Realism or reality may be used in several ways, i.e., a realistic outlook may imply that the individual looks at himself realistically or that he looks at the world as others do. For this study, as did Norman [13], we chose the latter aspect, a comparison of the individual's estimate of others with the group's estimate of others. The term as used here has been called insight by others. Steinmetz [16], in an objective study of insight, has defined it as the ability to perceive the world as others do. Thus the term called reality in the present study could be called insight since there is a comparison of how the individual's perception of others conforms to the group's perception of others. However, it has been shown [1, 6, 7] that this is only one aspect of insight, for in addition to being able to look at the world as others do, insight means being able to look at oneself objectively. Thus it was decided that a more adequate term for the present measure was reality. Operationally in this study, then, reality occurs when the individual states others have the trait in question and others actually do have it when the group is speaking of others. Specifically, the subject in answering the second form (about others), says they possess a certain trait; the majority of the group (51% or more) when taking the second form, also say that others possess the trait. The subject is then being realistic, i.e., he is seeing others as others see others — there is a conformity between the subject's and the group's estimate of others. He is given a point for reality each time this occurs.

For both total empathy and reality scores, all the questions for factors A, I, and N were used, since it is possible to empathize and be realistic on both desirable and undesirable traits. Also it was possible to obtain a score for both empathy and reality on both "Yes" and "No" answers, which was not possible with projection.

In testing the first hypothesis, it was necessary to beware of the fact that spuriously high correlations could result from intratrait comparisons. As noted above, projection scores were derived partially from the first form of the test when the group as a whole took the test answering for themselves. Empathy scores were also partially derived from this measure. Therefore, high intratrait correlations might result because both projection and empathy scores were derived from the same base. For example, high relationships might occur when relating A empathy to A projection, I empathy with I projection, etc. Hence, intertrait relationships must be examined, correlating every variable with every other. Therefore, A projection was correlated with I projection, I reality, I empathy, N projection, etc. Then, if significant correlations occurred between any variable derived from one trait and that derived from another, a further check on the first hypothesis could be offered. In addition, although this was not stated as a hypothesis, it could be empirically determined whether, if an individual projected in one trait, he would project in another, or if he were realistic in one trait, he would be realistic in another, etc.

To examine the first hypothesis, then, scores in projection, empathy, and reality were intercorrelated for both intertrait and intratrait comparisons. To test the second hypothesis, that reality and empathy would be positively correlated with adjustment, correlations were derived between scores in reality and empathy for traits A, I, and N, and actual test scores[1]

[1]These scores, of course, do not represent completely the scores the subject would have made had he been allowed the "?" alternative in answering the questionnaire the first time.

obtained for these traits. The actual scores were the operational measures of adjustment. The hypothesis was first entertained that projection would correlate negatively with adjustment. This was not entirely tenable because it is obvious that in scoring a test like GAMIN, when a subject denies undesirable traits in himself, he obtains a better "adjustment" score. This is correct, because all intratrait correlations, as shown later in Table 3, have a positive r between projection and adjustment, a result which may be an artifact of the scoring of the test.

Results

Tables 1 and 2 present correlations testing the first hypothesis. Table 1 shows that all intratrait correlations are significant at the 1% level in the predicted direction. The median of the three intratrait correlations (excluding total scores) between projection and reality is —.45; the median of the 3 r's between projection and empathy is —.65; and the median of the 3 r's between reality and empathy is .83. Using total scores, these medians are respectively —.41, —.65, and .77. Regarding intratrait correlations, then, it may be seen that the first hypothesis is readily supported. However, as noted above, these r's may have an element of spuriousness to them. Table 2, which shows intertrait correlations, presents the following data. There are six correlations between projection and reality, with a median of —.16, which is not significant. Only one of these six correlations is statistically significant, but all are in the direction predicted. Between projection and empathy, there are also six correlations, with a median of —.29, significant at the 5% level. Five of these six correlations are statistically significant, and all six are in the direction predicted. Between reality and empathy, there are again six correlations, the

Table 1

Intratrait Correlations among Experimental Variables

Relationship	r
A projection vs. A reality	—.35**
A projection vs. A empathy	—.41**
A reality vs. A empathy	.74**
I projection vs. I reality	—.45**
I projection vs. I empathy	—.65**
I reality vs. I empathy	.83**
N projection vs. N reality	—.49**
N projection vs. N empathy	—.66**
N reality vs. N empathy	.89**
Total projection (A + I + N) vs. total reality (A + I + N)	—.41**
Total projection (A + I + N) vs. total empathy (A + I + N)	—.65**
Total reality (A + I + N) vs. total empathy (A + I + N)	.77**

** Significant at the 1% level.

Table 2

Intertrait Correlations among Experimental Variables

Relationship	r
A projection vs.	
I projection	.50**
I reality	—.03
I empathy	—.23*
N projection	.41**
N reality	—.17
N empathy	—.22
A reality vs.	
I projection	—.17
I reality	.62**
I empathy	.30**
N projection	—.14
N reality	.26*
N empathy	.25*
A empathy vs.	
I projection	—.35**
I reality	.30**
I empathy	.50**
N projection	—.23*
N reality	.16
N empathy	.26*
I projection vs.	
N projection	.57**
N reality	—.29*
N empathy	—.37**
I reality vs.	
N projection	—.13
N reality	.48**
N empathy	.37**
I empathy vs.	
N projection	—.49**
N reality	.39**
N empathy	.52**

* Significant at the 5% level.
** Significant at the 1% level.

median of which is .30, significant at the 1% level. Again, five of these six *r*'s are statistically significant, and all are in the direction predicted by the hypothesis. Of the total of 18 intertrait *r*'s, all or 100% are in the indicated direction; 11 or 61% are statistically significant either at the 1% or 5% level, a figure which far exceeds chance expectations. The first hypothesis is thus supported by data from both Tables 1 and 2.

Each of the variables also correlates well with itself, according to Table 2. The median of the three correlations of projection in one trait with projection in another is .50, significant at the 1% level. The median of the three *r*'s of reality with reality is .48; of the three *r*'s of empathy with empathy is .50. Both these *r*'s are significant also at the 1% level. Thus it may be stated that if an individual projects in one GAMIN trait, he will project in another, if he is realistic in one he will be realistic in another, etc.

Table 3

Correlations between Adjustment and Experimental Variables

Relationship	*r*
A adjustment vs.	
A projection	.43**
A reality	.09
A empathy	.27*
I projection	.11
I reality	.24*
I empathy	.29*
N projection	.00
N reality	.00
N empathy	.06
I adjustment vs.	
A projection	.24*
A reality	.25*
A empathy	.33**
I projection	.13
I reality	.44**
I empathy	.56**
N projection	−.01
N reality	.22
N empathy	.31**
N adjustment vs.	
A projection	.11
A reality	.33**
A empathy	.42**
I projection	−.02
I reality	.38**
I empathy	.43**
N projection	.13
N reality	.21
N empathy	.38**
Total adjustment (AIN) vs.	
Total projection (AIN)	.18
Total reality (AIN)	.46**
Total empathy (AIN)	.56**

* Significant at the 5% level.
** Significant at the 1% level.

The second hypothesis stated that reality and empathy will correlate more closely with adjustment than will projection. Data bearing on this hypothesis are presented in Table 3. In this table, the median intratrait *r* of the three correlations between projection and adjustment is .13 (that all three are positive may be an artifact of scoring the GAMIN, as stated before). The median of the three intratrait *r*'s between reality and adjustment is .21; the median of the three between empathy and adjustment is .38. Only the latter median figure is statistically significant, but the order is in the direction of our hypothesis. Projection and adjustment give one rather high statistically significant *r* of .43, for trait A. Since all of our subjects were males, it is perhaps expected that denial of lack of ascendance would be important. Reality also has one statistically significant intratrait *r* with adjustment, but empathy correlates significantly in all three instances. Using total scores, the *r* of total adjustment on traits AIN with total projection on these traits is .18, with total reality it is .46, with total empathy it is .56. The latter two are significant at the 1% level.

If we examine the intertrait *r*'s which are free of contamination, it will be noted that the median *r* between projection and adjustment is just about .05 for the six correlations. Only one is statistically significant. The median *r* for the relationship between reality and adjustment is .25, significant at the 5% level. Among these six correlations, four are statistically significant. The median of the six *r*'s between empathy and adjustment is .32, significant at the 1% level. Five of these six *r*'s are statistically significant.

Thus, for both intra- and intertrait relationships, there appears a hierarchy of order of correlation with adjustment as follows: pro-

jection lowest, reality next, and empathy highest. From the standpoint only of significant *r*'s with adjustment, projection has two, reality five, and empathy eight, according to Table 3. The second hypothesis, then, appears to be supported.

Discussion

Our hypotheses are generally upheld by the data, the method used in the present study giving valid results supporting current personality theory. Certain methodological objections may be raised, especially that of a lack of knowledge about the reliability coefficient of the judgments of others (i.e., when the test was taken the second time), the fact that adjustment scores were changed in some unknown quantity by the elimination of the "?" alternative, the comparatively low *N* which made a "51% criterion" only 38 cases, and the assumption that self-reports were valid measures of adjustment. The latter criticism has been particularly mentioned by Ellis [9]. However, despite these objections, it is the writers' belief that this operational approach to the investigation of personality dynamics can be quite fruitful. It is certainly no accident that despite some of the crudities of the present investigation, *all* the correlations in Tables 1 and 2 were in the expected direction, and the trend of the correlations in Table 3, whether intra- or intertrait, was in the nature hypothesized at the outset. Further research might well employ outside criteria of adjustment, projection, empathy, and reality, to test whether our operational definitions have any validity.

Summary

An attempt was made, using operational definitions of terms, to verify two hypotheses. The first was that projection would be negatively related to both reality and empathy, but that the latter two variables would be positively related to each other. The second was that reality and empathy would correlate more closely with adjustment than would projection. Both these hypotheses were generally verified, using a personality questionnaire, the GAMIN, given under conditions wherein the subject took it for himself first, then made judgments about others. Consistent intra- and intertrait correlations in the direction predicted supported the first hypothesis. The second was supported by correlations giving a hierarchy of relation to adjustment from greatest to least of empathy, reality, and projection.

Received May 1, 1953.

References

1. Alexander, F. *Fundamentals of psychoanalysis.* New York: Norton, 1948.
2. Allport, G. W. *Personality, a psychological interpretation.* New York: Holt, 1938
3. Bellak, L. The concept of projection, an experimental investigation and study of the concept. *Psychiatry,* 1944, **7,** 353-370.
4. Cottrell, L. S., & Dymond, Rosalind F. The empathic responses. *Psychiatry,* 1949, **12,** 355-359.
5. DeCillis, Olga E., & Orbison, W. D. A comparison of the Terman-Miles M-F Test and the *Mf* scale of the MMPI. *J. appl. Psychol.,* 1950, **34,** 338-342.
6. Dymond, Rosalind F. The relation of insight and empathy. *J. consult. Psychol.,* 1948, **12,** 228-233.
7. Dymond, Rosalind F. A scale for the measurement of empathic ability. *J. consult. Psychol.,* 1949, **13,** 127-133.
8. Dymond, Rosalind F. Personality and empathy. *J. consult. Psychol.,* 1950, **14,** 343-350.
9. Ellis, A. The validity of personality questionnaires. *Psychol. Bull.,* 1946, **43,** 385-440.
10. Grossman, D. The construction and validation of two insight inventories. *J. consult. Psychol.,* 1951, **15,** 109-114.
11. Guilford, J. P., & Martin, H. G. *The Guilford-Martin inventory of factors GAMIN.* Beverly Hills, Calif.: Sheridan Supply Co., 1943.
12. Healy, W., Bronner, Augusta F., & Bowers, A. M. *The structure and meaning of psychoanalysis.* New York: Knopf, 1945.
13. Norman, R. D. The interrelationships among acceptance–rejection, self–other identity, insight into self, and realistic perception of others. *J. soc. Psychol.,* 1953, **37,** 205-235.
14. Remmers, H. H. A quantitative index of social-psychological empathy. *Amer. J. Orthopsychiat.,* 1950, **20,** 161-165.
15. Sears, R. R. Experimental studies of projection: I. Attribution of traits. *J. soc. Psychol.,* 1936, **7,** 151-163.
16. Steinmetz, H. C. Directive psychotherapy: measuring psychological understanding. *J. clin. Psychol.,* 1945, **1,** 331-335.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Moral Indignation: A Correlate of Denied Dislike of Parents' Traits

Sidney M. Jourard
Emory University

The hypothesis under consideration in the present study is stated as follows: *S*s who will admit to a predominant degree only favorable attitudes toward many traits of their parents will express a higher degree of moral indignation than *S*s who will express some attitudes of dislike toward their parents' traits.

Moral indignation is defined as the expression of annoyance, anger, or hostility toward transgressors of social mores.

The rationale for the hypothesis may be stated as repression-displacement theory of hostility. According to this theory, hostility toward themselves is inevitably generated by parents in their efforts to socialize their children. This hostility in the children may be allowed or denied expression. If it is denied, the children will learn to repress it, and displace it to hate objects provided by a culture for the guilt-free expression of hostility. Further, where hostility toward parents has been repressed, the children will tend to perceive the parents in undifferentiated, uncritical, almost magical terms, as Frenkel-Brunswik [1, pp. 340-346] found to be the case in highly authoritarian subjects whom she interviewed.

Procedure

Materials. Three questionnaires were devised to measure the relevant attitudes, viz: a moral indignation scale (MI), a scale to measure the feelings of each *S* toward 61 traits of his father (the Father-cathexis scale, C_F), and an identical scale to measure feelings toward the mother (Mother-cathexis scale, C_M).

Moral indignation scale.[1] Fifty statements were selected which seemed to describe "breaches of common decency" and which would stimulate moral indignation in a person. Many of these were paraphrased from items used in the F scale [1] and from Cason's study of annoyances [2], while others were devised by the writer.[2] The *S*s were required to indicate on the following five-point scale the degree of annoyance provoked by each situation:

1. Not annoyed
2. Slight irritation
3. Annoying
4. More annoying
5. Makes me furious

Sample items are as follows: A communist making a speech to you about his beliefs. A person who avoids the army draft because of "nervous" illness. The rapidity with which Hollywood stars change spouses.

Father-cathexis and Mother-cathexis scales.[2] Sixty-one traits were listed, and *S* was required to signify the nature of his feelings toward each of these traits, as he knew them in his father and mother.

The scale was devised as follows:

1. Have strong positive feelings: encircle a 1 for those aspects of the person which you like very much.

[1]The questionnaire forms have been deposited with the American Documentation Institute. Order Document No. 4032 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting $1.25 for 35 mm. microfilm or $1.25 for 6 x 8 in. photocopies. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

[2]Many ideas for items were obtained by perusal of letters to the editor of the local newspaper where manners and morals are commented upon, often with moral indignation. Finally, one important source of ideas for items deserves mention — the topics about which the writer, his colleagues, and many housewives gossip when gathered for social purposes.

2. Moderate positive feelings: encircle a 2 for those traits which you like in the person, but not as strongly as in category 1.

3. Have no feelings one way or the other: encircle a 3 for those aspects of the person about which you have no feeling.

4. Have moderate negative feelings: encircle a 4 for those aspects of the person which you dislike moderately.

5. Have strong negative feelings: encircle a 5 for those aspects of the person which you strongly dislike.

Sample items from the questionnaires are: sense of humor, temper, general knowledge, ability to express sympathy.

Subjects. One hundred and fifteen *S*s served in the study. These included 69 male college undergraduates, 24 college females, 17 females from a YWCA social club, and 5 Air Force men tested by a student assisting the writer. The age range was from 18 to 42, with a mean of 23 and an *SD* of 4.23.

Testing procedure. All the questionnaires were administered at one sitting—the MI first, and then the C_P and C_M scales. Instructions were as follows:

On these scales, you are asked to indicate your feelings. Please fill them out with as much self-honesty as you can. To assist you in this, we are asking you only to put down your birth date, so you need not fear what anyone might think concerning how you respond. Remember, please, to examine your feelings, so that your answers reflect these as accurately as possible.

Results

Means, standard deviations, range of scores, and odd-even reliability coefficients for the three tests are shown in Table 1.

All three tests appear to have high reliability and a broad range of scores.

The general hypothesis may be restated as follows: there will be a negative correlation between MI scores and both C_P and C_M scores.

The moral indignation scores were found to correlate −.39 with Father-cathexis, and −.37 with Mother-cathexis. Both *r*'s, while low, are significant beyond the .001 level, and support the hypothesis.

Discussion

The present data signify that persons who admit predominantly favorable attitudes to their parents tend to express a high degree of moral indignation, as defined and measured herein. Persons able to admit more critical feelings (higher C_P and C_M scores) toward their parents tend to produce lower MI scores.

It must be noted, however, that the correlations, despite their high confidence level, are low; a very small amount of the variance in MI scores is accounted for by variation in the parent-cathexis scores. This should not be surprising since there might be many reasons for the expression of moral indignation besides the displacement of repressed resentment of parents.

Table 1

Means, Standard Deviations, Ranges, and Reliability Coefficients for the Three Questionnaires

	Mean	σ	Range	Odd-even reliability coefficients*
Moral indignation	136.36	22.90	86–192	.90
Mother–cathexis	137.48	29.50	80–209	.95
Father–cathexis	136.38	28.08	67–206	.95

* Correlations are between odd and even halves of tests and are corrected by the Spearman-Brown formula.

Summary

One hundred and fifteen persons were tested with scales designed to measure moral indignation and attitudes toward parents. It was hypothesized that *S*s who admit little criticism and dislike of their parents would show higher indignation scores than *S*s who do admit some dislike of their parents' traits. Low but highly significant correlations between MI scores and C_M and C_P scores supported this hypothesis.

Received May 18, 1953.

References

1. Adorno, T. W., Frenkel-Brunswik, Else, Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Harper, 1950.
2. Cason, H. Common annoyances, a psychological study of every-day aversions and irritations. *Psychol. Monogr.*, 1930, 40, No. 2 (Whole No. 182).

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

The Comparability of Measures of Profile Similarity: An Empirical Study

James N. Mosel
The George Washington University
and June B. Roberts
Pennsylvania State College

In recent years there has been an increased emphasis on the use of psychometric profiles in clinical diagnosis, vocational guidance, personnel evaluation, and even social psychology. This increased emphasis has resulted largely from the realization that interpretation of an individual's set of scores, for diagnosis or prediction, must frequently be based upon the pattern or configuration of scores, rather than upon consideration of one test at a time or upon an additive sum of scores. Computation of mean score differences between nosological or criterion groups, or the derivation of linear combinations of test scores to predict a criterion (as in the multiple-regression equation and discriminant function), exemplify the latter, nonconfigural approach. It has been felt by many, especially clinicians, that these conventional techniques are inadequate because they fail to take into account the fact that the diagnostic or predictive meaning (i.e., weight) of a score on one test is not constant, but varies according to the scores obtained on the other tests. Consequently, the argument goes, interpretation must consider patterns of scores. This is mathematically tantamount to saying that interpretation involves a nonlinear combination of scores—nonlinear in that the weights assigned to each test are not constant but vary as a function of score on the test itself and on the other tests. It is this set of intertest relationships as defined by a nonlinear combination that is implied by such terms as "pattern," "configuration," "profile effect," etc. Discussion of the problems and concepts of configural vs. linear treatments of multiple measurements will be found in Horst [6] and Mosier [9].

The great interest in patterns has quite naturally led to the problem of matching profiles—measuring the degree of similarity in the pattern of two profiles. The usual situation in clinical practice in which this need arises is where a standard or reference profile has been established[1] and it is desired to match a given "unknown" or referred profile against it. The assumption underlying this is that the closeness of the match measures the extent to which the unknown profile will display certain criterion or nosological characteristics associated with the standard profile.

The need for profile matching also arises in the research situation when it is desired to learn whether the profiles of one diagnostic category are more similar to each other than they are to the profiles of some other diagnostic category. The principle here is that if the within-groups similarity is not greater than between-groups similarity, the groups cannot be distinguished by profile patterns. This is illustrated in a study by Kogan [7], who attempted to discriminate between schizophrenic and neurotic patients by means of Wechsler subtest patterns. In a similar vein, *Q* technique attempts to determine whether the agreements among the profiles of persons are such that when factor analyzed, a typology can be established. Investigations of this type, incidentally, provide one approach to the clinical problem

[1]The problem of identifying and specifying standard profiles is beyond the purview of this paper, but it is quite apparent that systematic methods are yet to be developed. It should be noted that the conventional procedure of computing the mean scores of a given criterion group and then using the resulting profile of averages as the standard completely ignores the possibility that the number of characteristic profiles may be greater than one.

of establishing the standard or reference profile mentioned above. Before a standard profile can be established, it must first be shown that a basic profile exists which is typical of the criterion group and which differentiates this group from others.

There are two aspects of profiles on which matching may occur: the *shape* or configuration of scores and the general *elevation* (also called intensity or level) of scores as indicated by some measure of central tendency. Thus we may speak of matching for *absolute* agreement, in which we consider similarity of both shape and elevation, or for *relative* agreement, in which we consider only similarity of shape.

Considerable discussion in the literature [2, 3] has been devoted to the relative merits of measuring similarity in terms of shape and elevation or of shape alone. The basic question is whether profiles of similar shape but of different elevations indicate the same diagnosis or prediction. Much of the confusion apparently stems from the failure to realize that there is no single, best conception of profile similarity and that the decision between the two alternatives rests on what particular characteristics interest the investigator and on what empirical studies show he can obtain. Profiles can yield several types of information, some of which are perhaps more clearly related to the shape of the profile, while others are more closely associated with both shape and elevation. In many instances the ultimate answer can be given only by empirical research designed to show which approach provides the greatest differentiation with respect to an external criterion reflecting the investigator's purpose.

One interesting variation of the empirical approach is that of Barnette [1], who evaluated separately differences in shape and elevation of aptitude profiles. That this approach may be fruitful is further supported by Schmidt's work [14] with the MMPI. He found that groups diagnosed as severe and mild psychoneurosis had highly similar profile shapes but different elevations, the severe group having a more elevated profile than did the mild.

There are, of course, certain situations where no outside criterion is available or even meaningful, such as in Osgood's use of profiles to display connotative meaning [10]. In such cases an entirely rational decision is not only necessary but proper.

Indices of Profile Similarity

In an attempt to overcome the obvious disadvantages of matching profiles by visual judgment, a number of statistical indices of profile similarity have been proposed. These are described below, but because the present study is concerned with the empirical agreement among these, no attempt will be made to elaborate upon or evaluate their statistical properties.[2]

Correlation coefficient. Perhaps the most obvious and generally employed statistic is the product-moment or rank-difference correlation coefficient. The correlation technique yields a nonarbitrary measure which reflects similarity of shape but disregards elevation.[3] This index ranges from 1.00 for high similarity, through zero for no similarity, to − 1.00 for high inverse similarity (i.e, when the two profiles are mirror images of one another). The correlation coefficient has a disadvantage in that profiles of rather dissimilar shape can have equal correlation with the standard profile. Also, the value of 1.00 does not necessarily indicate perfect similarity; two profiles can be quite dissimilar in shape and still correlate perfectly.

Meehl's difference index. This measure [8] reflects shape but neglects elevation. The computational procedure is as follows:

1. Obtain for each profile the deviations (keeping algebraic sign) of the scores of each test from the second highest test in the profile. These values may be symbolized as d.
2. Obtain the differences between the corresponding deviations in the two profiles. These values may be symbolized as δ.
3. Square the δ's after first rounding to the nearest tenth.[4]
4. Sum the squared δ's of the N tests and extract the square root.

In formula form:

$$I = \sqrt{\sum_{i}^{N} \delta^2_{jl}}$$

where $\delta_{jl} = d_{ij} - d_{il}$; and $d_{ij} = X_{ij} - S_j$ and $d_{il} = X_{il} - S_l$, in which X_{ij} and X_{il} are corresponding scores of profiles j and l on test i (of N tests), and S_j and S_l are the second highest scores in profiles j and l respectively.

Meehl states that his index tells how far

[2] For a general discussion of various indices, see Cattell [2], Du Mas [3], and Gaier and Lee [5].

[3] Intraclass correlation, however, considers both shape and elevation differences [16]. This is because the deviations are taken from the grand mean of both sets of profile scores. On the other hand, this index suffers from a restricted negative range.

[4] Meehl [8] reports that the rounded values correlate .95 with the unrounded.

apart two profiles are in the hyperspace of N dimensions in which each profile is projected as a point. The smaller the value the closer the match, although the index is an arbitrary one, having no absolute meaning. Since its value depends on the number of tests, it is comparable only for matchings in which all pairs of profiles have the same number of tests. Meehl reports that the index correlates .68, .65, and .91 with subjective estimates of similarity made visually by three judges. This index is very new and Meehl admits that it is frankly experimental and that its statistical properties remain to be determined.

D *coefficient.* Osgood and Suci [11] have described an index designed to reflect both shape and elevation. Although these authors apparently consider their proposal as new, the same method was suggested some time ago by Sagebeer [13] and more recently by Toops [15], who calls it the "precise profile method." The D coefficient is based on the geometric principle that in a space of N mutually orthogonal dimensions, the distance between two points is equal to the square root of the sum of the squared differences between the coordinates of the points on each dimension. Since profiles may be considered as points in N space (where N equals the number of tests), the distance between them serves as a measure of similarity. This distance is D, which in formula form is:

$$D = \sqrt{\sum_{i}^{N} d^2_{jl}}$$

where $d_{jl} = X_{ij} - X_{il}$, in which X_{ij} and X_{il} are the corresponding scores of profiles j and l on test i of N tests. The computational procedure is simple:

1. Obtain the difference between corresponding scores on each test for the two profiles.
2. Square and sum these differences.
3. Extract the square root.

Like Meehl's index, the D coefficient is an arbitrary number whose value depends on the number of tests in the profile. It should also be noted that the above formula applies only when the tests are uncorrelated.

Du Mas' coefficient of profile similarity. Du Mas [4] has introduced the coefficient of profile similarity, r_{ps}, which describes similarity in terms of the agreement in slope between corresponding segments of the two profiles being matched. By "segment" is meant the line drawn from one score point on the psychograph to another adjacent point. The slope is positive when the left score point of a segment is closer to the bottom horizontal of the graph than is the right score point. Zero slope occurs when the segment is parallel to the bottom horizontal.

In formula form:

$$r_{ps} = 2\ (S/T - .50),$$

where S is the sum of the corresponding profile segments which exhibit slope of the same sign, and T is the total number of segments in which no zero slope occurs. The standard error of r_{ps} is given by Du Mas as

$$\sigma_{r_{ps}} = \sqrt{1/T},$$

thus enabling a test of the null hypothesis. Du Mas [4] states that the assumptions underlying the coefficient are that the test score distributions are symmetric and identical, the regression of one test on another adjacent test is linear, and the sequence of the tests remains fixed.

Obviously, r_{ps} considers only profile shape and, in addition, loses precision because it is based only upon the sign of the slope, not the magnitude. Consequently, it is quite possible, as Du Mas himself points out, to obtain a very high index of similarity when the two profiles are quite dissimilar in both shape and elevation. While r_{ps} can range from -1.00 to 1.00, it cannot, however, always assume a zero value. Moreover, it yields a discrete rather than continuous distribution of values. The number of possible values it can assume can never be greater than $N + 1$ (where N is the number of tests in the profile). For a four-test profile r_{ps} can assume only five values: 1.00, 0.50, 0.00, $-$ 0.50, and $-$ 1.00. Also, the fact that the index is an arbitrary function of the order in which the tests are arranged may constitute a criticism. This feature is not shared by the other indices described here. The index has the advantage, however, of being rapid to compute, especially in view of the table Du Mas gives for cases where $T \leq 30$.

Cattell's coefficient of pattern similarity. This coefficient, r_p, is one of three indices developed by Cattell [2] to measure different aspects of profile similarity, the other two being his coefficient of pattern shape, r_s, and his coefficient of pattern-effect, r'_p. He concluded

that the most generally applicable is r_p since this index is designed to reveal both profile shape and elevation.

The present paper will therefore be confined to r_p which is defined by the formula

$$r_p = \frac{2k - \Sigma d^2}{2k + \Sigma d^2}$$

where k is the median for chi square (i.e., the value of chi square at $p = .50$) with degrees of freedom equal to $N - 1$, where N is the number of tests in the profile; and d is the difference between corresponding scores on the two profiles at each test, expressed in standard score form with respect to the distribution within that test.

Cattell states that for practical purposes the index can be considered as approximately similar to an ordinary correlation coefficient. It ranges from 1.00, meaning perfect similarity, through zero, meaning complete lack of relation, to −1.00, meaning complete inverse relationship. The assumptions are that the tests are independent and that the distribution of each test is normal.

Purpose of the Present Investigation

In view of the above multiplicity of measures, a crucial question relates to the amount of agreement these measures would display in rank ordering a set of profiles in similarity to a standard profile. The purpose of the present study was to determine empirically the comparability of these measures as revealed by their intercorrelations. Inasmuch as these measures were originally devised to replace the presumably less reliable method of subjective matching, an additional purpose was to compare the statistical indices with the results from visual judgment. The question also arises as to whether the agreement among indices varies according to the standards used. Correlations were therefore computed separately for matchings in which differing profiles were used as standards.

Procedure

Actual profiles were used, consisting of scores on the first five subscales of the Kuder Preference Record as obtained from a group of supervisors in a government agency. Four profiles were selected as standards because of their extreme visual dissimilarity. A group of 30 other profiles served as "referred" profiles. Using the five indices described above, measures of similarity to each standard were computed for each referred profile.

Subjective Matching

In measuring similarity by visual judgment, an effort was made to secure reliable judgments. Apparently no studies have been made of the reliability of judges or of methods in rating the similarity of profiles; consequently, several preliminary studies were made using two judges and three rating methods. Judge A was a psychologist; judge B, an engineer. For whatever significance it may have, both judges were above the 85th percentile on the Minnesota Paper Formboard.

Method of single stimuli. The judges rated each profile against each standard on the following six-point scale: 1: "almost identical," 2: "very similar," 3: "fairly similar," 4: "somewhat similar," 5: "no resemblance," 6: "completely opposite." To permit the judge to establish a stable subjective scale of judgment, all profiles were matched against one standard before rating against the next standard was begun. Each judge repeated his ratings one week later.

Rank-order method. Next, each profile was ranked in order of similarity to each standard. The judges first rank ordered the profiles previously assigned to each rating-scale interval. When profiles had thus been ranked within each group, they could readily be placed in an over-all rank order of similarity to the standard.

Method of triads. As a further check on the reliability of the above rank orders, judge A assigned a second set of rank-order values using the method of triads [12]. In this method the judge compares *two* referred profiles with the standard and decides which of the two most resembles the standard. As in the method of paired comparisons, all possible pairs of referred profiles are thus compared with the standard. The total number of times each profile is judged most similar can then be used to yield rank orders.

Results

Judgmental matching. Table 1 shows the rate-rerate reliabilities of the two judges from the method of single stimuli (correlations ex-

pressed as Pearson r) and from the method of rank order (correlations expressed as rho). Table 2 shows the interjudge agreement for the two methods, basing the intercorrelations on the mean of each judge's first and second ratings.

Table 1

Rate-Rerate Reliabilities of Two Judges in Matching 30 Profiles against Four Standard Profiles by the Method of Single Stimuli and the Method of Rank Order

Standard	Judge A		Judge B	
	Single stimuli*	Rank order	Single stimuli	Rank order
1	.75	.87	.85	.87
2	.91	.95	.88	.90
3	.92	.95	.70	.78
4	.83	.76	.46	.25

* The correlations for the method of single stimuli are given in Pearson r; those for the method of rank order in rho.

Table 2

Intercorrelation between Two Judges in Matching 30 Profiles against Four Standards by the Method of Single Stimuli and the Method of Rank Order

Standard	Single stimuli (r)	Rank order (ρ)
1	.28	.53
2	.80	.85
3	.50	.67
4	.71	.58

It will be noted that the rank-order method tends to yield higher reliabilities than does the method of single stimuli. As might be expected, the interjudge reliabilities are consistently lower than the rate-rerate reliabilities.

These results indicate that both judges are fairly self-consistent but differ somewhat in their definitions of "similarity." Judge A has a very acceptable reliability, being somewhat superior to judge B. Judge A's reliability is further evidenced by a rho of .80 between the mean of his first and second values from the rank-order method and the method of triads. This correlation is almost as great as A's rate-rerate reliability by the rank-order method. The conclusion is that the ratings of judge A by the rank-order method are the most reliable. These were therefore retained for comparison with the various statistical indices of profile similarity.

Index agreement. Table 3 shows the matrix of intercorrelations (rank difference) among the five statistical indices of profile similarity and the subjective judgments of judge A. The correlations have been determined separately for each standard profile.

Table 3

Intercorrelations among Five Statistical Measures and Visual Judgments of Profile Similarity; 30 Profiles Matched against Four Standard Profiles

(Correlations expressed as rho)

Measure	Meehl	Osgood	Du Mas	Cattell	Pearson
Standard 1					
Osgood	.73				
Du Mas	.005	.30			
Cattell	.81	.51	.04		
Pearson	.70	.28	.17	.60	
Judge A	.65	.41	.50	.78	.49
Standard 2					
Osgood	.81				
Du Mas	.82	.73			
Cattell	.65	.64	.73		
Pearson	.80	.71	.74	.54	
Judge A	.80	.77	.90	.68	.75
Standard 3					
Osgood	.60				
Du Mas	.49	.46			
Cattell	.32	.31	.48		
Pearson	.14	.42	.20	−.006	
Judge A	.29	.28	.57	.55	−.002
Standard 4					
Osgood	.65				
Du Mas	.25	−.11			
Cattell	.59	.48	.12		
Pearson	.54	.40	.23	.50	
Judge A	.38	.06	.40	.47	.17

It is obvious that the type of standard employed makes a sizable difference in the amount of agreement among the measures. For instance, the Du Mas and Meehl indices correlate .005 when the profiles are matched against Standard 1, but .82 when Standard 2 is used. This means, of course, that the extent to which these measures are tapping a common factor of "similarity" varies according to the profile adopted as the standard. The comparability of the measures cannot, therefore, be stated in absolute terms; it depends on the standard against which the profiles are matched.

Inspection of Table 3 reveals that the two measures which disagree most with the other measures are the Du Mas coefficient[5] and Pearson *r*, especially the latter. Cattell's coefficient agrees most closely with subjective judgments of similarity, while the indices of Du Mas, Meehl, Osgood, and Pearson follow in that order. The indices of Meehl and Osgood intercorrelate more highly on the whole than do any other pairs of measures. This is somewhat surprising in view of the fact that the former reflects only profile shape, while the latter reflects both shape and elevation.

Summary

The agreement among five statistical indices of profile similarity and visual judgments was studied through intercorrelation. A group of 30 profiles from the Kuder Preference Record was matched against four different standard profiles, and the resemblance was measured by Cattell's coefficient of pattern similarity, Du Mas' coefficient of profile similarity, Osgood and Suci's *D* coefficient, Meehl's index, and Pearson *r*. Visual judgments of profile similarity by two judges using two methods (single stimuli and rank order) showed acceptable rate-rerate reliabilities; the interjudge reliabilities were somewhat lower. The scale values from the most reliable judge using the most reliable method (rank order) were adopted for comparison with the results from the five statistical indices.

Results showed a wide degree of disagreement among the measures as to the rank order in which they arranged a set of profiles with respect to similarity to the standard profile, the rank-difference intercorrelations ranging from −.006 to .90. Also, the amount of agreement between any two measures of profile similarity varied considerably according to the particular profile employed as the standard, thus indicating that the comparability of the measures cannot be stated in absolute terms but is contingent upon the standard used.

Cattell's coefficient showed the closest agreement with visual judgments, while the indices of Du Mas, Meehl, Osgood and Suci, and Pearson followed in that order. The greatest agreement among the statistical indices was that between the coefficients of Meehl and Osgood.

[5]The correlation of Du Mas' r_{ns} with the other indices is undoubtedly reduced because of the restricted range of values which this index can assume. No correction for widespread classes was made, however, because it was desired to ascertain the interrelationships in terms of the categories given.

Received April 1, 1953.

References

1. Barnette, W. L. Occupational aptitude patterns of selected groups of counseled veterans. *Psychol. Monogr.*, 1951, **65**, No. 5 (Whole No. 322).
2. Cattell, R. B. r_p and other coefficients of pattern similarity. *Psychometrika*, 1949, **14**, 279-298.
3. Du Mas, F. M. On the interpretation of personality profiles. *J. clin. Psychol.*, 1947, **3**, 57-65.
4. Du Mas, F. M. The ciefficient of profile similarity. *J. clin. Psychol.*, 1949, **5**, 123-131.
5. Gaier, E. L., & Lee, Marilyn. Pattern analysis: the configurational approach to predictive measurement. *Psychol. Bull.*, 1953, **50**, 140-148.
6. Horst, P. The prediction of personal adjustment. New York: *Soc. Sci. Res. Coun. Bull.*, 1941, No. 48.
7. Kogan, W. S. An investigation into the relationship between psychometric patterns and psychiatric diagnosis. *J. gen. Psychol.*, 1950, **43**, 17-46.
8. Meehl, P. E. *Minnesota Multiphasic Personality Inventory research for counselors.* St. Paul: St. Paul Dept. of Educ., 1950.
9. Mosier, C. I. Batteries and profiles. In E. F. Lindquist (Ed.), *Educational measurement.* Washington, D. C.: Amer. Coun. on Educ., 1951. Pp. 764-808.
10. Osgood, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, **49**, 197-237.
11. Osgood, C. E., & Suci, G. J. A measure of relation determined by both mean difference and profile information. *Psychol. Bull.*, 1952, **49**, 251-262.
12. Richardson, M. W. Multi-dimensional psychophysics. *Psychol. Bull.*, 1938, **35**, 659. (Abstract)
13. Sagebeer, R. Identifying patterns of ability. *Educ. Rec. Suppl.*, 1938, **19**, 172-178.
14. Schmidt, H. O. Test profiles as a diagnostic aid: the MMPI. *J. appl. Psychol.*, 1945, **29**, 115-131.
15. Toops, H. A. Philosophy and practice of personnel selection. *Educ. psychol. Measmt*, 1945, **5**, 95-124.
16. Webster, H. A note on profile similarity. *Psychol. Bull.*, 1952, **49**, 538-539.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Children's Responses to Animal and Human Figures in Thematic Pictures[1]

Mary Ann Smith Armstrong
Institute of Child Welfare, University of Minnesota[2]

The Children's Apperception Test (CAT) [1] is a projective test for children primarily between the ages of three to ten years. The CAT is based on the theory that children can identify more readily with animal figures than with human figures.

The problem of this study was to determine which picture series elicits more thematic material. The hypotheses of the writer are as follows:

1. Pictures which contain human figures will evoke longer fantasy productions.
2. More verbalization which is nondescriptive in content will occur in protocols based on human figures.
3. Nouns, verbs, and ego-related words will have higher frequencies in stories having human figures as their main characters.
4. Responses to pictures depicting human relationships will involve more "blocking," which, in turn, will produce a longer reaction time.

Method

The subjects were 60 children who attended the first, second, and third grades in the University of Minnesota elementary school. The sample included 10 boys and 10 girls from each grade. The mean age of the first grade subjects was 6-7, with a range from 6-2 to 7-3. Second grade subjects ranged in age from 7-3 to 8-2, with a mean age of 7-6. Children in third grade had a mean age of 8-6, with a range from 8-1 to 8-11. The mean IQ of first grade subjects measured by the 1937 Revision of the Stanford-Binet was 120, with a range from 76 to 156. Children in the second grade had a mean IQ of 131, with a range from 108 to 174. Third grade subjects ranged from 97 to 172, with a mean IQ of 123. The subjects were predominately of upper socioeconomic status.

Bellak's CAT pictures were ranked by 10 judges on the basis of their supposed stimulus value in evoking verbalizations from children. Graduate students and faculty served as raters. Five pictures from the original CAT were selected for use in this study on the basis of the mean ratings obtained. The experimental pictures (Series B) depicted a feeding situation (Card 1), attitudes toward parents (Card 2), sibling rivalry (Card 4), acceptance by adult figures (Card 8), and conceptions of punishment (Card 10). A duplicate set of pictures corresponding to the CAT in style and composition was drawn by Heidi Schwabacher, graduate art student at the University of Minnesota. The duplicate pictures (Series A) contained human figures in place of animal figures; in all other respects the two sets of pictures were equivalent. Both picture series were ambiguous with respect to sex.

In order to counteract the facilitating effects of the initial test on the final test with respect to increased productivity, two groups of 30 children were chosen from the original sample by means of random numbers. Each group consisted of an equal number of boys and girls from each grade, thereby equalizing the factor of sex differences within the group. The two sets of pictures were administered in random order, using a test-retest method: Group A was given Series A, followed by Series B; Group B received Series B, followed by Series A. Retests were administered following a time interval of two weeks.

[1]The writer wishes to express her appreciation to Dr. Dale B. Harris for his helpful suggestions in the preparation of this manuscript.

[2]Now with State Department of Education, St. Paul.

Instructions to the subjects were held constant for both sets of pictures. The experimenter presented each picture to the child with the request, "Tell me a story about this picture." To clarify a point in the child's story the experimenter said, "Tell me more," "What do you mean?" or "Explain it to me." At appropriate places in the story the experimenter asked, "What happened before?" and "What will happen later?" The experimenter was limited to three questions in the administration of a single picture.

Measurement of the subject's reaction time was held constant for each picture. The child's response was timed with a stop watch from the presentation of the picture to the completion of the story. Total reaction time was recorded in seconds for each protocol.

Analysis of the fantasy data according to length of the response was done using the scoring system developed by McCarthy [2] in an earlier language study. Contractions of the subject and predicate counted as two words. Hyphenated words and compound nouns referring to single objects scored as one word. Contractions of the negative and the verb counted as a single word. Each part of the verb form was scored as one word.

Indices of the nouns, verbs, and ego-related words were computed on the basis of their frequency in the individual protocol. A word counted as a noun or verb depending on its function in the sentence. Compound nouns and verb combinations were considered single words. The phrases, "it looks like," "it seems as though," "it seems like," and the personal pronouns, I, my, me, and mine were scored in the Ego-Related Word Index. The response "I don't know" was not classified as an ego-related word.

Analysis of the stories in terms of nondescriptive verbalization was done using the transcendence categories outlined by Weisskopf [3]. The Transcendence Index of a picture was obtained by counting the number of comments about the picture which exceeded pure description. The mean number of nondescriptive remarks per subject constituted the index of the picture.

Two additional categories were used in computing a Transcendence Index, a sex-reference category and a naming category. The category of sex reference refers to the characterization of a figure in relation to its sex by the use of a singular or plural word. The naming category is the characterization of a figure by the use of proper nouns. The terms "baby" and "child" were not scored in any category.

The total number of seconds required by the subject to project himself into a particular picture was considered an index of reaction time.

Results

Means and standard deviations were obtained for length scores by sex and grade (Table 1). With grade progression, children tend to tell longer stories. Girls give slightly but consistently longer protocols than boys. There is no consistent tendency for animal pictures to evoke longer protocols than human pictures.

An analysis of nouns and verbs used, by the procedure employed in Table 1, showed slight but expected grade and sex differences, but no

Table 1

Length of Response Means and Standard Deviations by Grade and Sex for Human and Animal Pictures

Grade	Picture	Boys		Girls		Total	
		M	*SD*	M	*SD*	M	*SD*
First	Human	61.48	29.25	64.54	36.45	63.01	33.08
	Animal	75.22	36.45	56.54	36.49	65.88	37.65
Second	Human	66.64	33.13	68.66	52.13	67.65	43.69
	Animal	64.38	33.02	70.80	35.77	67.59	34.57
Third	Human	62.30	44.33	86.02	46.09	74.16	46.75
	Animal	65.12	38.58	88.28	58.17	76.70	50.70

picture differences; the data are omitted in the interest of brevity.

Table 2 presents the means and standard deviations for ego-related word scores by grade and sex. A developmental trend may be disclosed in the grade means. The use of first person pronouns shows a slight but consistent decline with grade progression. Again, there are no differences attributable to picture source.

Table 2

Ego-Related Word Means and Standard Deviations by Grade and Sex for Human and Animal Pictures

Grade	Picture	Boys		Girls		Total	
		M	*SD*	M	*SD*	M	*SD*
First	Human	1.92	1.41	2.64	1.73	2.28	1.73
	Animal	2.86	2.00	2.04	1.41	2.45	1.73
Second	Human	1.24	1.00	1.86	1.41	1.55	1.41
	Animal	1.64	1.73	1.62	1.73	1.63	1.73
Third	Human	1.62	1.41	1.46	1.41	1.54	1.41
	Animal	1.50	1.73	1.54	2.23	1.52	2.00

Mean transcendence scores and standard deviations by sex and grade are found in Table 3. The means for the three grades suggest a difference between the two series of pictures. Children give a greater number of nondescriptive comments in their responses to human figures than to animal figures. Girls tend to make more statements which are subjective and personalized than boys.

Table 3

Transcendence Means and Standard Deviations by Grade and Sex for Human and Animal Pictures

Grade	Picture	Boys		Girls		Total	
		M	*SD*	M	*SD*	M	*SD*
First	Human	7.98	3.74	9.00	5.91	8.49	4.89
	Animal	7.42	3.74	6.86	4.79	7.14	4.24
Second	Human	9.30	3.74	9.50	6.16	9.40	5.09
	Animal	6.86	2.64	7.58	4.35	7.22	3.60
Third	Human	8.32	3.87	10.20	5.09	9.26	4.58
	Animal	7.12	4.24	9.50	6.24	8.31	5.47

Means and standard deviations were computed for reaction time scores by grade and sex. The data, omitted in the interest of brevity, show that girls have a consistently longer reaction time than boys in all grades.

Analysis of variance procedures were applied to the various scores reported in Tables 1 to 3 and in the text above. Table 4 summarizes the results of these various analyses. The sex differences conform to the well-known superiority of girls in various language measures. The grade differences, too, are not surprising. The results with respect to human and animal pictures are significant. Mean lengths of protocols and mean reaction times for the two sets do not differ. Nor do the sets differ in the number of naming responses, action words, or ego-related words. There is a significant difference between the animal and human figures with respect to Weisskopf's transcendence scores. Contrary to the assumption of the Children's Apperception Test, the difference favors the human pictures, which evoke significantly more subjective, personalized, interpretive responses, other than pure description. It is to be recalled that these conclusions are based on normal elementary school children, somewhat superior in intelligence. On every measure the variance among the means of children was significantly greater than the average variance within a child's responses to the several pictures.

Table 4

Results of Analyses of Variance Applied to Different Scores Derived from Thematic Picture Protocols

[Entries indicate the level of significance satisfied by the *F* ratio of the between-means variance tested against the remainder variance (the within-children variance)]

Comparison	Between means of sexes	Between means of grades	Between means of pictures	Between means of children
Length of protocol	.05	.01		.01
Number of nouns	.01	.01		.01
Number of verbs	.05			.01
Number of ego words		.01		.01
Transcendence scores	.01		.01	.01
Reaction time	.05			.01

Discussion

The original hypothesis that human pictures will evoke no more thematic material than animal pictures is refuted. The hypothesis that human pictures will evoke more nouns, verbs, and ego-related words fails to be refuted, as does the hypothesis that human pictures will induce more "blocking" as measured by reaction time to the picture.

Summary

Two duplicate series of pictures, one using animal and the other using human figures, were given to 60 children according to the following plan. Two groups of 30 children were chosen from the original sample by means of random numbers. Each group consisted of an equal number of boys and girls. Group A was tested with Picture Series A, followed by Picture Series B; Group B received Picture Series B, followed by Picture Series A. The picture series were separated by a time interval of two weeks. The arrangement of the pictures in a series for each subject was randomized by recourse to a table of random numbers.

1. With grade progression, children produce longer protocols. As children advance in grade level, they employ more nouns and fewer first person pronouns in their protocols.

2. Girls use a greater number of nouns and verbs in their protocols than boys. Girls employ more words and more nondescriptive comments in their protocols than boys. Girls have a slightly longer reaction time than boys.

3. Human figures evoke more verbalization which goes beyond pure description than do animal figures.

4. All scores used in this study are subject to greater variation among children than average variation "within children" (i.e., between all pictures for one child).

Received May 25, 1953.

References

1. Bellak, L. *Children's Apperception Test.* New York (P.O. Box 42, Gracie Station): C.P.S. Co., 1949.
2. McCarthy, Dorothea A. *The language development of the preschool child.* Minneapolis: Univer. of Minnesota Press, 1930.
3. Weisskopf, Edith A. A transcendence index as a proposed measure in the TAT. *J. Psychol.,* 1950, 29, 379-390.

Journal of Consulting Psychology
Vol. 18, No. 1, 1954

Delinquent vs. Nondelinquent Performance on the Porteus Qualitative Maze Test [1]

Richard F. Docter and C. L. Winder
Stanford University

Porteus [4] has described a "qualitative" method for scoring the Maze Test which appears to reflect differences in test behavior of delinquents and nondelinquents. This scoring scheme is based on an evaluation of the subject's regard for instructions, carefulness, impulsivity, etc. Essentially, it is a measure of expressive movements, and differs from the scoring of the mazes for intelligence ratings.

Studies by Porteus [2, 3], Wright [5], and Grajales [1] have revealed very significant differences between the qualitative performances of delinquents and nondelinquents. However, only Porteus has tested nondelinquents, the other authors having accepted his nondelinquent norms. (See Table 1.)

This study was conducted for the following two reasons: (*a*) to provide additional nondelinquent normative data, and (*b*) to provide an independent study of "matched" groups of delinquents and nondelinquents.

Procedure

Groups of 60 delinquent and 60 nondelinquent boys, matched for age, mental ability, race, and socioeconomic level, were given the Maze Test. The delinquents comprising this group were chosen for their similarity to those studied in previous investigations. All subjects were Caucasians, and were between the ages of 13 years, 10 months and 16 years, 10 months. Wechsler-Bellevue intelligence ratings for the delinquents ranged from 72 to 116. All delinquents were inmates in a state school for delinquent boys.

[1]The writers express their gratitude to Robert Van Vorst and the staff of the Fred C. Nelles School for Boys at Whittier, California, and to personnel of the Public School System of San Jose, California.

Nondelinquents were selected from the ninth grade of a junior high school and the tenth grade of a high school in San Jose, California. Their intelligence as measured by the California Test of Mental Maturity ranged from 71 to 128. In order to exclude noninstitutionalized delinquents from the nondelinquent group, the names of all subjects were reviewed by the school counselor and by the boys' physical education teachers. Only those who survived this process of elimination were tested. While the delinquent and nondelinquent groups are by no means perfectly matched, it appears certain that the most striking difference between groups is that one is characterized by "good citizenship," while the other is distinguished by its members' persistent disregard for the rules of society.

Results

The mean *Q* score (total weighted qualitative score) for delinquents was 47, which does not differ significantly from 49, the central tendency of previously reported means for delinquents (Table 1).

Table 1
Q-Score Means and Sigmas for Groups of Delinquents and Nondelinquents

Group	Delinquents			Nondelinquents		
	N	Mean	Sigma	*N*	Mean	Sigma
Porteus [2]	100	49	25	100	22	13
Wright [5]	54	49	28	—	—	—
Grajales [1]	60	56	29	—	—	—
Porteus [3]	50	48	22	179	19	13
Present study	60	47	26	60	25	20

The mean *Q* score for nondelinquents was 25, as compared to 22 which Porteus reported. This difference of three *Q*-score points is not significant. The major finding, then, is that the

qualitative performance of delinquents and nondelinquents corresponds almost exactly with the results previously reported for such groups, and that the difference between the delinquents' and nondelinquents' means yields a critical ratio in excess of five, which is significant at the .0001 level.

Table 2

Percentages of Delinquents and Nondelinquents Falling above Various *Q*-Score Critical Cutoffs

Group	Delinquents		Nondelinquents	
	Cutoff	Per cent above	Cutoff	Per cent above
Porteus [2]	29	80	29	21
Wright [5]	29	78	—	—
Grajales [1]	22	92	—	—
Present study	29	70	29	30

It is interesting to consider what predictive value this test might have. Porteus suggests a critical cutoff of 29 as the *Q* score which best discriminates delinquents from nondelinquents (Table 2). The writers found that of the delinquents, 70% fell above 29, while only 30% of the nondelinquents scored above this level. On the basis of these consistent differences between groups, this test shows promise of having predictive value.

Porteus probably constructed his weighting scheme after observing differences between the mean incidence of weighted errors for the two groups. Table 3 permits comparison of the means given by Porteus with the results of the present study, and it is noteworthy that the means reported by Grajales [1] and Wright [5] are very similar to these.

In order to assess the appropriateness of the Porteus weights, chi squares were computed based on differences in error incidence between the two groups. All the qualitative errors are discriminating in the expected direction (delinquents higher than nondelinquents), and on the basis of the magnitude of the respective chi squares, the Porteus weights appear to be reasonably assigned. Certain modifications could be justified on the basis of these data, but such changes would be premature.

That Porteus' weights are not essential is suggested by two findings. First, the critical ratio between the unweighted error means is 5.18, which is approximately the same as the critical ratio between the *Q*-score means. Second, using an unweighted error cutoff of 16, almost exactly the same percentages of each group fall above and below as when a *Q*-score cutoff of 29 is employed. It appears that the Porteus weights serve primarily to increase the magnitude of scores without altering the differences between groups. In the present study, the unweighted scores were: (*a*) delinquent mean = 24.7, $\sigma = 13.2$; (*b*) nondelinquent mean = 13.7, $\sigma = 9.9$.

Unfortunately, there are no data available on the reliability of this measure, but this study revealed a very high correlation between two scorers. Using the records of 60 nondelinquents, the product-moment correlation was .98; the rank-difference correlation, .96. Had the records of delinquents been used, these correlations might have dropped as the delinquents committed more errors which were difficult to score objectively; on the other hand, the increased range for the delinquents could compensate for this.

Table 3

Mean Incidence of Qualitative Errors, the Porteus Weights, and Chi Squares for Differences between Delinquents' and Nondelinquents' Performance

Error	Mean incidence of qualitative errors				Porteus weights	Chi square	*p*
	Nondelinquents		Delinquents				
	Porteus [2]	Present study	Porteus [2]	Present study			
First third	1.01	0.60	1.58	1.63	2	4.76	.05
Last third	0.77	0.60	1.09	0.80	1	1.20	.30
Cut corner	3.01	3.15	4.40	4.61	1	7.00	.01
Cross line	3.18	3.50	5.93	7.68	2	13.04	.001
Lift pencil	1.63	2.80	5.77	3.81	3	2.40	.20
Wavy line	1.34	1.10	3.68	4.36	2	40.14	.001
Wrong direction	0.55	0.33	0.71	0.83	1	2.30	.20
Year V	0.27	0.28	1.09	0.46	1	2.84	.20
Year VI	0.33	0.21	0.99	0.50	1	2.48	.20

The extent to which Q score correlates with intelligence is a vital question. The results of the study corroborate previous reports which indicate the correlation is low. For delinquents who had taken the Wechsler-Bellevue, the r was $-.27$, significant at the .05 level. The nondelinquents' Q score correlated $-.16$ with the California Test of Mental Maturity, not significant. These should be interpreted as positive correlations, since a low Q score is "good."

In this research the entire series of mazes was given each subject, beginning with year V and continuing through the adult maze. The Q score was derived from this sequence, but the Maze Test intelligence rating was obtained only from that portion of this series which was applicable.

Since some subjects require two trials on the easier mazes and three or four trials to solve the more difficult ones, it might be asked what influence the number of mazes administered has upon the Q score. The range of mazes used was sufficient to justify the computation of a product-moment correlation between this variable and Q score. Combining both groups, the r was .29, which reaches the .01 level of significance. Therefore only about 9 per cent of the variance of Q score is attributable to variations in number of mazes used. It is recommended that the entire series, including the adult maze, be given.

Discussion

If one accepts the empirical demonstration that this test psychometrically distinguishes groups of adolescent boys whose social behavior is known to be different, there remains the question: What does the test measure? We confess that a justifiable answer for this question is not available at this time. But one guess which the writers suggest as reasonable is that this measurement of expressive movements taps one or more personality variables related to the construct of ego-control. Impulsivity may best describe the cluster of responses which contribute to the Q score. Should this be the case, and future research may clarify the matter, instruments of this kind would appear to be an important contribution toward the objective assessment of personality.

Summary

1. Groups of 60 delinquents and 60 nondelinquents were given the Porteus Qualitative Maze Test. The mean total weighted qualitative score (Q score) for delinquents was 47; for nondelinquents, 25. The difference was significant at the .0001 level.

2. Seventy per cent of the delinquents scored above 29; only 30 per cent of the nondelinquents exceeded this cutoff. The test shows promise of having predictive value.

3. Porteus' weighting scheme can be rationalized and used, but does not seem to add to the efficiency of the test.

4. A correlation of .98 showed reliability of scoring can be very high.

5. For delinquents, Q score correlated $-.27$ with intelligence, significant at the .05 level. For nondelinquents the correlation was $-.16$, not significant.

6. Q score correlated .29 with the number of mazes used, significant at the .01 level.

7. This test may measure personality variables, such as impulsivity, related to the construct of ego-control.

Received April 6, 1953.

References

1. Grajales, M. C. Porteus' Qualitative Maze Test as a measure of delinquency. Unpublished master's thesis, Fordham Univer., 1945.
2. Porteus, S. D. Qualitative performance in the maze test. Vineland, N. J.: Smith Printing House, 1942.
3. Porteus, S. D. Q-scores, temperament and delinquency. *J. abnorm. soc. Psychol.*, 1945, **21**, 81-103.
4. Porteus, S. D. *The Porteus Maze Test and intelligence.* Palo Alto, Calif.: Pacific Books, 1950.
5. Wright, Clare. The qualitative performance of delinquent boys on the Porteus Maze Test. *J. consult. Psychol.*, 1944, **8**, 24-26.

NEW BOOKS AND TESTS

Books

Brown, Judson S., Harlow, Harry F., Postman, Leo J., Nowlis, Vincent, Newcomb, Theodore M., & Mowrer, O. Hobart. *Current theory and research in motivation: a symposium.* Lincoln: Univer. of Nebraska Press, 1953. Pp. v + 193 (paper). $2.00.

You will want to read this one. The six able contributors to the symposium are at their best and are, in the order named, as predictably Hullian, experimental, perceptual, developmental, social, and two-factored as you might expect. Would the subject matter were as predictable as the protagonists! But they shed more light on the complex problem of motivation than has been shed in many a moon. If you are interested in psychological theory, send to Nebraska for your copy — now. — *L.F.S.*

Daniel, Robert S., & Louttit, C. M. *Professional problems in psychology.* New York: Prentice-Hall, 1953. Pp. xi + 416. $5.50.

This book supplies a sound core around which to build a professional problems course for graduate students in the applied fields of psychology. It contains a good assortment of materials relevant to the skills which psychologists must have in addition to their subject matter — about jobs, training, history, bibliography, written communication, publication, organizations, ethics, legislation, and the like. The style is clear and graceful but mature; it does not sidestep complex and difficult issues by talking down to the reader. The graduate student, very soon to be a member of the profession, is recognized as interested in and capable of understanding the affairs of his field. The book integrates well with basic source materials such as the APA's *Publication Manual* and *Ethical Standards of Psychologists.* It is in agreement with these documents and supplements them without undesirable duplication. Now that this good resource is available there should be more and better orientation courses for our students. And many already-graduated psychologists may read it with enjoyment and profit, too. — *L.F.S.*

Note. — The reviews were prepared by the Editor and the Associate Editors, who may be identified by their initials.

Doll, Edgar A. *The measurement of social competence: a manual for the Vineland Social Maturity Scale.* Philadelphia: Educational Test Bureau, 1953. Pp. xviii + 664. $7.75.

The Vineland Social Maturity Scale is an unusual psychological instrument, not a test but a method for quantifying observations of social maturation. It is one of the few devices that can be applied reliably and usefully in the absence of the examinee from the reports of persons who have observed his developmental achievements. The Scale, available since 1936, is now provided with a comprehensive manual which enhances its usefulness. The manual describes the conception and construction of the scale, its administration and scoring, and its standardization and validation. The author draws both from his own basic studies and from the published research of others. Most of the evidence supports the value of the Scale: its good repeat reliability, its effective scaling of items, and its applicability to many practical and research problems. It is something of a shock, however, to find that the present standardization is based on only 20 cases, 10 male and 10 female, at each age from 0-1 to 30. The paucity of cases is offset by the unusual care exercised in selecting the normative subjects and by the supporting data from consecutive ages. The manual, like the Scale, has many merits and a few clear shortcomings. The completeness of the data presented, and the cautiousness of the interpretations, may serve as models for other test authors. On the negative side, the manual, while attractive in appearance, is organized so as to make essential data hard to find; the index seems to provide frustration as often as help. The diligent searcher, however, is ultimately rewarded. The Social Maturity Scale has now come to its own maturity, and should become as essential as a clinical tool as the intelligence test. — *L.F.S.*

Dollard, John, Auld, Frank Jr., & White, Alice Marsden. *Steps in psychotherapy.* New York: Macmillan, 1953. Pp. ix + 222. $3.50.

This is an exciting little book. In simple language, the authors undertake to orient the reader to the Dollard-Miller theory of therapy, "a blend of reinforcement learning theory and psychoanalysis," and then proceed to demonstrate the implications of the theory for practice by presenting selected portions of the seventeen recorded therapy sessions constitut-

ing brief psychotherapy for a 40-year-old married woman patient. Discussion of these materials is largely in terms of the recorded comments of the therapy supervisor (Dollard) to the young therapist (an anonymous psychiatrist) during the supervisory sessions. Part II includes the complete protocols of the battery of tests administered between the first and second therapy hours and two independent sets of test interpretations by clinical psychologists, one written with and the other without access to a summary of the first therapy session. The authors do not claim validity for either the therapeutic or diagnostic procedures but rather offer these raw materials and their ideas as promising approaches to the development of better knowledge. The book should be extremely useful as a basis for seminar discussions and of interest to all psychologists concerned with the training of psychotherapists. — *E.L.K.*

Eysenck, H. J. *The structure of human personality.* London: Methuen, and New York: Wiley, 1953. Pp. xix + 348. $5.75.

Eysenck's most recent contribution to the study of personality is a survey of research. But it is more than a survey; it is a critical and creative integration which brings order to a diversity of quantitative studies. Eysenck espouses factor analysis as the prime technique for revealing the organization of personality. In an introductory chapter on theory he identifies first-order factors with "traits" and second-order factors with "types." The research evidence is reviewed in a series of chapters arranged according to the technique of appraisal employed: ratings, questionnaires, objective tests, measurements of physique, physiological indices, attitude and interest measurements, and correlations between persons. Many studies are reinterpreted, sometimes with new analyses of their data, so as to bring a semblance of agreement to results previously considered contradictory. Eysenck does not indulge in the petty virtue of impartiality but the blows he gives his opponents, while often severe, are fair and well aimed. The book will appeal to psychologists and students who prefer theory buttressed by research to vague professions of faith. They may not agree with every word but they will learn and be stimulated. — *L.F.S.*

Festinger, Leon, & Katz, Daniel. (Eds.) *Research methods in the behavioral sciences.* New York: Dryden, 1953. Pp. ix +660. $5.90.

No cookbook of investigative techniques, this stimulating volume represents a serious and vigorous attempt to clarify the logic of a number of research problems in the sciences of individual and collective human behavior. Such methods as the sample survey, the field study, and the formal experiment under both field and laboratory conditions are analyzed in terms of their advantages and disadvantages with respect to various substantive issues in social and psychological theory. Two chapters are devoted to the technical problems of sample selection and distribution-free statistical procedures, both of which should prove especially meaningful for research workers in clinical psychology and in the psychology of personality. One section deals with techniques of observing group behavior, of using documents and census records, and of collecting usable data by interview. On the analytic side, there are contributions on the use of qualitative data and on the theory and methods of social measurement. While it is a little unfair in a book of such remarkably even and high quality to give special mention to particular sections, Helen Peak's chapter on the general problem of making objective observations could profitably be read by every graduate student in the country. Similarly, the concluding chapter by Likert and Lippitt on the utilization of social science is an insightful discussion, at once seminal and sound, of the neglected problem of how research results can be most fruitfully capitalized upon both by other researchers and by applied workers and citizen leaders. This book will richly repay serious and repeated consideration by psychologists of very different theoretical persuasions and specific investigative interests. — *E.J.S.*

Hathaway, Starke R., & Monachesi, Elio D. (Eds.) *Analyzing and predicting juvenile delinquency with the MMPI.* Minneapolis: Univer. of Minnesota Press, 1953. Pp. viii + 153. $3.50.

A research monograph containing six studies based on the complete testing of a ninth-grade group with the MMPI in 1947-48, and a subsequent comparison with police and court records. Moderate relationships between some MMPI indicators and delinquency were found which can be interpreted with the aid of expectancy tables. The most valid prediction was the nondelinquency of youngsters with clearly normal profiles. A valuable side product of the study is a useful amount of data about adolescents in general. — *L.F.S.*

Kinsey, Alfred C., Pomeroy, Wardell B., Martin, Clyde E., & Gebhard, Paul H. *Sexual behavior in the human female.* Philadelphia: Saunders, 1953. Pp. xxx + 842. $8.00.

There will probably be as many lines of type set about this book as there are in it, which should make a reviewer cautious lest he review its stimulus value rather than the volume per se. The book is similar in many respects to the 1948 report on the male. It differs in that it does not contain the large part on "factors affecting"; instead, a part is devoted to a comparison of the sexes. The Kinsey team admits it has not undertaken "U. S. corrections" of the raw data. However, there is a point-for-point comparison with male results, which raises questions about the significance of these generalizations. The authors frankly admit the inad-

equacy of their sample, especially for females who have not gone beyond grade school. Female sexual behavior is studied with orgasm and erotic arousal as the dependent variables, with age, education, marital status, occupation, religious and rural-urban backgrounds, age of onset of adolescence, and decade of birth as independent variables. In the male study mere sexual arousal was not included as a variable. Religious background and decade of birth were conspicuously important variables; age, a very important variable in male outlet, was related differently to female behavior. Males were found to be conditioned by sexual experience more frequently than were females, and they share vicariously in the sexual experiences of others to a greater extent. The data for youth show majority behavior following somewhat the course dictated by cultural pressures. Kinsey and his associates did more than report findings. They suggested social implications which go beyond the scope of their molecular study. Whereas the authors generally seemed to show scientific caution and scholarliness, some of the earlier criticism of the limitations of their method does not seem to have influenced the second study greatly. — *F. McK.*

Lehman, Harvey C. *Age and achievement.* Princeton, N. J.: Princeton Univer. Press, 1953. Pp. xiii + 359. $7.50.

Psychologists will be grateful for this volume which brings together for the first time Lehman's penetrating studies of age and achievement, previously published in journal articles over an interval of nearly twenty years. Like many good empiricists, Lehman raises more questions than he answers. Why is outstanding creativity achieved earlier in chemistry than in the social sciences? Why do our creative scientists seem to be getting younger, while our statesmen are getting older? These facts, and many others as interesting, are attested by Lehman's careful studies. The "whys" remain vital topics for additional investigation. — *L.F.S.*

Peck, Leigh. *Child psychology.* Boston: Heath, 1953. Pp. 536. $5.25.

Child Psychology is planned for college students on the undergraduate level. The author has constantly kept in mind the special needs of students preparing to work with children, and has given a comprehensive picture of psychological development. It is refreshing to find a book that makes its subject matter such delightful reading for the group for which it is planned. Pertinent reports of actual situations illustrate the discussions and help the reader clarify his own thinking. The arrangement, organization, references, and index are excellent and the book should prove helpful in an introductory course in child development. — *B.M.L.*

Piaget, Jean. *The child's conception of number.* New York: Humanities Press, 1952. Pp. ix + 248. $5.00.

This is a translation of a book previously published in French some twelve years ago. Piaget's naturalistic approach to the mind of the child is here applied to his development of the number function. As usual, Piaget has included material from the original protocols, and the book makes interesting reading for the psychologist. The practicing clinician will get more insight here into the functioning of the child's mind in a testing situation than he will from the abstractions of factor analysis. The translation is good and the format pleasing. — *W.A.H.*

Powdermaker, Florence B., & Frank, Jerome D. *Group psychotherapy.* Cambridge: Harvard Univer. Press, 1953. Pp. xv + 615. $6.50.

This is a detailed report of two experiments in group therapy. Both studies were conducted under the auspices of the Washington School of Psychiatry, one with 124 VA clinic patients, mostly neurotics, and the other with 88 VA hospital patients, mostly schizophrenics. Only for the psychotic group was a control group used. The magnitude of the experiment can be seen from the fact that 24 "medical doctors" were involved (10 of them psychiatrists), 5 clinical psychologists, and 2 psychiatric social workers. There were 12 groups in the clinic experiment and 15 in the hospital. In the neurotic group, 453 sessions were studied. The major method employed was the situational analysis of group meetings, developed from Harry Stack Sullivan's concept of situations as bases of human behavior. The analyses make excellent teaching (and learning) material. An equally valuable portion is an appendix with a running account of the first and the 65th meeting of one group. Situations are also presented to illustrate processes in the conduct of group therapy. One aspect of these studies will probably evoke some controversy. Fourteen of the 24 M.D.'s apparently had had no previous clinical training, but they were permitted to do "psychotherapy" with their groups. In fact, according to the authors, their lack of experience turned out to be an advantage. On the other hand, the clinical psychologists and psychiatric social workers were not employed in the group therapy. One is tempted to ask some questions about the reasoning involved in these decisions. However, the importance of this volume lies in the discussions of research methods in group psychotherapy, rather than in the results obtained by the therapy. — *M.K.*

Sargent, Helen D. *The Insight Test.* Menninger Clin. Monogr. Ser. No. 10. New York: Grune & Stratton, 1953. Pp. xii + 276. $6.75.

Sargent has supplied a more complete manual for her ingenious pencil-and-paper projective method

which was first described in 1944. In the test, subjects' written responses to described conflict situations are scored for expressed affect (*A*), defense (*D*), and malignancy (*M*) and for several subscores under each of these main categories. Much use is made of the ratio *A/D*. The manual gives the theory underlying the test, its administration, scoring, and interpretation, and copious illustrative protocols drawn from seven clinical groups. When judged by the standards of a quantitative test, the method receives meager support. Although the scoring reliability of trained scorers is good, the retest reliability of subjects is reported only for inadequate groups and even then ranges from poor to mediocre. The "normal" group, on which standards are based, consists of 20 screened university students. Obviously, evidence that patients differ from such "normals" has no value. The author gives evidence of being a competent statistician, and discusses the needs for controls of age, sex, education, and socioeconomic status but does nothing about them. A fair inference is that she does not really believe in experimental evaluation. The test, of course, can be judged by other criteria. The clinical protocols reveal it as a rich source of fertile hunches about psychopathology. Until better research evidence is produced, however, the *Insight Test* must remain a better device for making hypotheses than for testing them. — *L.F.S.*

Viteles, Morris S. *Motivation and morale in industry.* New York: Norton, 1953. Pp. xvi + 510. $9.50.

This volume, written primarily for management, "is devoted to a comprehensive description and critical evaluation of American and British experimental studies and attitude surveys bearing upon the sources of motivation and the determinants of morale in industry." It is indeed comprehensive and the bibliographies at the end of each chapter are extensive. Some background in motivational theories is given. The problem of mobilizing the will-to-work is discussed with particular reference to wage-incentive systems. Studies on the effect on productivity of such variables as job preference, size of work group, knowledge of results, supervision, and employee participation are covered effectively. There is also a section on employee attitude surveys and a final one on guideposts for management, with cautions on application of such surveys and the other techniques discussed. — *A.R.*

Tests

Finch, F. H., Kuhlmann, F., & Betts, G. L. *Kuhlmann-Finch Intelligence Tests.* Kgn.-gr. 12, 9 levels. 1 form, each level. 25 (30) min. Test booklet ($1.95 per 25) with instructions, keys; IBM answer sheet (95¢ per 25); manual, pp. v + 57 ($1.00); specimen set ($1.50). Philadelphia: Educational Test Bureau, 1953.

The new series of tests bears only a superficial resemblance to its predecessor, the Kuhlmann-Anderson. The tests consist of entirely new items planned by Kuhlmann and developed after his death by Finch. Each booklet has five subtests with lenient five-minute timing. Nonverbal tests are used through grade 3, and two verbal subtests, two nonverbal, and one numerical at the higher levels. MA and IQ scores are obtained from charts in the booklets, with little clerical effort. The manual gives only 10 pages to the construction and standardization of the test, and there are no references to research articles. While not extensive, the technical data give a decidedly favorable impression. Items were selected not only to yield progressive age increments, but to be as free as possible of cultural and sex biases. Reliabilities are good (.86 — .92); norms are based on large and widespread samples. Data on validities, except against the general developmental criterion, are still lacking. — *L.F.S.*

King, Joseph E., & Wingert, Judith W. *IPI Merit Rating Series.* Industrial employees. 5 scales, clerical, mechanical, sales, supervisor, technical. Rating form ($4.00 per 20); manuals, pp. 20, 8; specimen set ($3.00). Tucson, Ariz. (Box 6056): Industrial Psychology, Inc., 1953.

The *Merit Rating Series* applies a psychological test concept to the rating of employee performance. In place of the usual rating scale, 60 statements are provided in the form for each job family. The rater marks each statement *yes* or *no*, and weighted scoring yields quantitative scores for several performance traits and for over-all competence. The manuals discuss the methods of rating, training of raters, and application of results. At least minimally acceptable data are given on item difficulties and consistencies, the derivation of weights, and norms. In addition to supervisory uses, the method seems to be of potential value as a criterion for studies of personnel aptitude and selection. — *L.F.S.*

Books Received

Abrahams, Joseph, & Varon, Edith. *Maternal dependency and schizophrenia.* New York: International Universities Press, 1953. Pp. 240. $4.00.

Bonaparte, Marie. *Female sexuality.* New York: International Universities Press, 1953. Pp. x + 225. $4.50.

Cottle, Wm. C. *The MMPI: a review.* Lawrence, Kan.: Univer. of Kansas Press, 1953. Pp. vi + 82.

Donahue, Wilma, Rae, James Jr., & Berry, Roger B. (Eds.) *Rehabilitation of the older worker.* Ann Arbor: Univer. of Michigan Press, 1953. Pp. v + 200. $3.25.

Ginzberg, Eli, Herma, John L., & Ginsburg, Sol W. *Psychiatry and military manpower policy.* New York: King's Crown Press, 1953. Pp. xi + 66. $2.00.

Greenacre, Phyllis. (Ed.) *Affective disorders.* New York: International Universities Press, 1953. Pp. 212. $3.00.

Heston, Joseph C. *How to take a test.* Chicago: Science Research Associates, 1953, Pp. 47. 40¢.

Ingram, Christine P. *Education of the slow-learning child.* New York: Ronald, 1953. Pp. ix + 359. $5.00.

Kahn, Alfred. *A court for children.* New York: Columbia Univer. Press, 1953. Pp. xxii + 359. $4.50.

Leitner, Konradi. *Hypnotism for professionals.* New York: Stravon, 1953, Pp. 127. $4.00.

Little, Wilson, & Chapman, A. L. *Developmental guidance in secondary school.* New York: McGraw-Hill, 1953. Pp. ix + 324. $4.50.

Nichtenhauser, Adolf, Coleman, Marie L., & Ruhe, David S. *Films in psychiatry, psychology and mental health.* New York: Health Education Council, 1953. Pp. 269. $6.00.

Prosser, Charles A., & Sifferd, Calvin, S. *Selecting an occupation.* Bloomington, Ill.: McKnight & McKnight, 1953. Pp. x + 246. $2.50.

Ralli, Elaine P. (Ed.) *Adrenal cortex* (Transactions of the Fourth Conference, 1952). New York: Josiah Macy, Jr. Foundation, 1953. Pp. 165.

Reik, Theodor. *The haunting melody.* New York: Farrar, Straus, & Young, 1953. Pp. viii + 376. $4.00.

Sartre, Jean-Paul. *Existential psychoanalysis.* New York: Philosophical Library, 1953. Pp. viii + 275. $4.75.

Serog, Max. *New light on dreams.* Boston: House of Edinboro, 1953. Pp. 159. $3.00.

Tyler, Leona E. *The work of the counselor.* New York: Appleton-Century-Crofts, 1953. Pp. x + 323. $3.00.

Vernon, Philip E. *Personality tests and assessments.* London: Methuer, 1953. Pp. ix + 220. $4.00.

Whitaker, Carl A., & Malone, Thomas P. *The roots of psychotherapy.* New York: Blakiston, 1953. Pp. xvii + 236. $4.50.

Wortis, Joseph. (Ed.) *Basic problems in psychiatry.* New York: Grune & Stratton, 1953. Pp. vi + 186. $4.50.

Yost, Orin Ross. *What you should know about mental illness.* New York: Exposition Press, 1953. Pp. 165. $3.50.

Psychologists....
who use projective techniques

will be interested in the following

Psychological Monographs

No. 303—An Experimental Examination of the Thematic Apperception Technique in Clinical Diagnosis by A. A. Hartmann. Pp. 48. $1.00.

No. 309—Rorschach Responses Related to Vocational Interests and Job Satisfaction by Solis L. Kates. Pp. 34. $1.00.

No. 310—Symbol Elaboration Test (S.E.T.): The Reliability and Validity of a New Projective Technique by Johanna Krout. Pp. 67. $2.00.

No. 314—The Validity of a Multiple-Choice Projective Test in Psychopathological Screening by Martin Singer. Pp. 40. $1.00.

No. 315—A Normative Study of the Thematic Apperception Test by Leonard D. Eron. Pp. 48. $1.50.

No. 316—Experimentally Induced Variations in Rorschach Performance by Edith E. Lord. Pp. 34. $1.00.

No. 325—A Dual Approach to Rorschach Validation: A Methodological Study by James O. Palmer. Pp. 27. $1.00.

No. 331—A Psychological Study of Eminent Biologists by Anne Roe. Pp. 63. $2.00.

No. 334—Reliability and Validity of Judges' Ratings of Adjustment on the Rorschach by Marguerite Q. Grant, Virginia Ives, and Jane H. Ranzoni.

No. 337—The Validity of Personality-Trait Ratings Based on Projective Techniques by Henry Samuels. Pp. 22. $1.00.

No. 338—Rorschach Estimates of Personality Attributes in the Michigan Assessment Project by Woodrow W. Morris. Pp. 31. $1.00.

ORDER FROM

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth Street N. W. Washington 6, D. C.

NEUROLOGY & PSYCHIATRY

Section VIII of Excerpta Medica

The Neurology Section includes abstracts of all articles on neurological surgery in addition to general clinical neuro-physiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition, the following special branches are dealt with fully:

* **Psychoanalysis, individual and analytical psychology**
* **Sexology, criminology, alcoholism, and drug addiction as related to psychiatry**
* **Social and industrial psychology and psychiatry, vocational guidance and mental testing**
* **Psychometrics**
* **Personality testing and the Rorschach test**
* **Heredity and statistical studies as they apply to psychiatry**
* **Mental defect and epilepsy**

Annual volume of 1000 pages, including an index classified by author and subject.

Subscription price $22.50 per annual volume

EXCERPTA MEDICA SERVICE CORPORATION

250 Madison Ave. **New York 16, N.Y.**

Please enter my subscription to Section VIII, NEUROLOGY AND PSYCHIATRY, OF EXCERPTA MEDICA. I enclose $22.50.

Name_____

Address_____